Essays on Growth and Innovation Policies

by

Ufuk Akcigit

Bachelor of Arts in Economics, Koc University, Istanbul (2003)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of
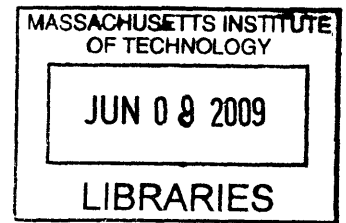
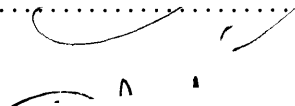Doctor of Philosophy

at the

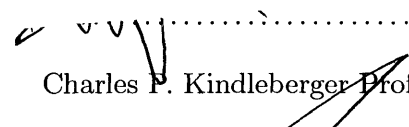## MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2009
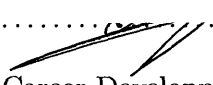
Signature of Author..
Department of Economics
May 15, 2009

Certified by.......
Daron Acemoglu
Charles P. Kindleberger Professor of Applied Economics
Thesis Supervisor

Certified by..
Mikhail Golosov
Dornbusch Career Development Associate Professor of Economics
Thesis Supervisor

Accepted by..........
Esther Duflo
Abdul Latif Jameel Professor of Poverty Alleviation and Development Economics
Chairman, Departmental Committee on Graduate Studies

Essays on Growth and Innovation Policies

by

Ufuk Akcigit

Submitted to the Department of Economics
on May 15, 2009, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis consists of three chapters on innovation and economic growth. Chapter 1 is a joint work with Daron Acemoglu. We study the form of intellectual property rights (IPR) policy and licensing in the context of endogenous economic growth with step-by-step cumulative innovation. The main questions of the analysis are as follows. Should a company with a large technological lead receive the same IPR protection as a company with a more limited advantage? Should technological followers be able to license the products of technological leaders? We propose a general equilibrium framework to investigate these questions. IPR policy regulates whether followers in an industry can copy the technology of the leader and also how much they have to pay to license past innovations. We prove the existence of a steady-state equilibrium and characterize some of its properties. We then quantitatively investigate the implications of different types of IPR policy on the equilibrium growth rate and welfare. The two major results of this exercise are as follows. First, the growth rate and welfare in the standard models used in the (growth) literature can be improved significantly by introducing a simple form of licensing. Second and more importantly, full patent protection is not optimal from the viewpoint of maximizing welfare; instead, welfare-maximizing (and growth-maximizing) policy involves state-dependent IPR protection, providing greater protection to technological leaders that are further ahead than those that are close to their followers. This form of the welfare-maximizing policy is a result of the "trickle-down" effect, which implies that providing greater protection to firms that are further ahead of their followers than a certain threshold increases the R&D incentives also for all technological leaders that are less advanced than this threshold.

Chapter 2 is an empirical study that analyses the relationship between firm size and innovation. The common practice in the endogenous growth literature is to assume a constant innovation quality (size) and focus only on innovation frequency. In this chapter, I test the validity of this assumption using Compustat and USPTO patent data to examine the relationship between firm size and innovation quality. Since R&D investment is the input and firm growth is the result of innovation, this chapter studies the relationship between firm size - firm growth and firm size - R&D intensity as well. The reduced form results uncover three stylized facts: Smaller firm grow faster, are more R&D intensive and more interestingly, produce higher quality innovations. These results are robust, among many other things, to sample selection and to differences in patenting behaviors.

In Chapter 3, I propose a theoretical model to understand the microfoundations underlying these stylized facts. In this model, technologically heterogenous firms compete for innovation.

A novelty of this model is that firms can endogenously choose not only the probability of innovation, but also the innovation quality. I prove the existence of the equilibrium and show that the model's predictions are consistent with the aforementioned reduced form evidences. These results rely on two assumptions: the concavity of the profit function with respect to firm size and the constant returns to scale property of the R&D production function. The intuition is that, when profits are concave, the incentives for radical innovation diminishes as firm size increases. As a result, R&D intensity and firm growth also decrease. Next, I estimate the structural parameters of the model using Simulated Methods of Moments and use the results to conduct a policy experiment. Since firms in this model do not internalize the positive externalities that they generate on other firms, there is underinvestment in R&D, so that there is scope for policy action through size-dependent R&D subsidies. In conclusion of this policy experiment, the optimal size-dependent R&D subsidy policy does considerably better than optimal uniform (size-independent) policy. Moreover, supporting small firms is more growth-enhancing than subsidizing big firms. For a large range of values for the elasticity of substitution, the growth enhancing effect dominates the negative impact of public spending on initial consumption. Therefore, the optimal (welfare-maximizing) policy provides higher subsidies to smaller firms.

Thesis Supervisor: Daron Acemoglu
Title: Charles P. Kindleberger Professor of Applied Economics

Thesis Supervisor: Mikhail Golosov
Title: Dornbusch Career Development Associate Professor of Economics

Acknowledgement

I am deeply indebted to my primary advisor Daron Acemoglu for my professional development. I have been very fortunate to learn from him as his student, research assistant and advisee during my PhD years. Our conversations have been the most rewarding moments of my life at MIT. He has constantly inspired me with his boundless enthusiasm for economics, his dedication to his students, his constant giving to others and his devotion to academia. My gratitude to him is infinite.

Special thanks to Mike Golosov and Philippe Aghion for sharing their experiences with me and providing support, guidance and constructive comments. Their enthusiastic advice was invaluable in my research.

The chapters of this dissertation benefited greatly from the discussions with George-Marios Angeletos, Nick Bloom, Wesley Cohen, Ulrich Doraszelski, Gino Gancia, Bronwyn Hall, William Kerr, Sam Kortum, Josh Lerner, Suzanne Scotchmer, Robert Shimer, Robert Townsend, and the seminar participants at the FED - Board of Governors, the Canadian Institute of Advanced Research, EPFL, European Science Days, Koc University, MIT, National Bureau of Economic Research Economic Growth and Productivity Groups, Technology Policy Conference and Toulouse Information Technology Network.

My dream of doing a PhD at MIT started during my undergraduate years at Koc University. My undergraduate advisor, Tayfun Sonmez, was the first person in my sophomore year to believe in me, and to encourage and support me to pursue my goal. He has been an enormous influence on my academic career for which I will always be grateful. I am also thankful to my undergrad professors Mete Soner, Insan Tunali, Utku Unver and Kamil Yilmaz who generously provided me their support at the beginning of this journey.

I never hesitated to lean on my friends, Gokhan Dogan, Ufuk Karabiyik, Pinar Kurt, Onur Mumcu and Terrance O'Neil, when I needed their help in difficult times. They always showed me how lucky I am to have them in my life.

Contents

Chapter 1

State-Dependent Intellectual Property Rights Policy

Joint with Daron Acemoglu

1.1 Introduction

How should the intellectual property rights of a company be protected? Should a firm with a large technological lead receive the same intellectual property rights (IPR) protection as a company with a more limited technological lead? These questions are central to many discussions of patent and competition policy. A recent ruling of the European Commission, for example, has required Microsoft to share secret information about its products with other software companies (New York Times, December 22, 2004). There is a similar debate about whether Apple should make iPod's code available to competitors that are producing complementary products. Central to these debates is the substantial technological lead that these companies have built over their rivals, which was viewed by the European Commission both as a source of excessive monopoly power and as an impediment to further technological progress in the industry. A systematic analysis of these policy questions and a full investigation of the effects of intellectual property rights on growth and welfare require a framework incorporating *state-dependent* patent/IPR protection policy. By state-dependent IPR policy, we mean a policy that makes the extent of patent or intellectual property rights protection conditional on the technology gap

between different firms in the industry. Existing work has investigated the optimal length and breadth of patents assuming an IPR policy that does not allow for licensing and is uniform. In this chapter, we make a first attempt to develop a framework that is rich enough to investigate these issues and we use it to study the implications, and the optimal form, of various IPR policies.

Our basic framework builds on and extends the step-by-step innovation models of Aghion, Harris and Vickers (1997) and Aghion, Harris, Howitt and Vickers (2001), where a number of (typically two) firms engage in price competition within an industry and undertake R&D in order to improve the quality of their product. The technology gap between the firms determines the extent of the monopoly power of the leader, and hence the price markups and profits. The purpose of R&D by the follower is to catch up and surpass the leader (as in standard Schumpeterian models of innovation, e.g., Reinganum, 1981, 1985, Aghion and Howitt, 1992, Grossman and Helpman, 1991), while the purpose of R&D by the leader is to escape the competition of the follower and increase its markup and profits. As in racing-type models in general (e.g., Harris and Vickers, 1985, 1987, Budd, Harris and Vickers, 1993), a large gap between the leader and the follower discourages R&D by both. Consequently, overall R&D and technological progress are greater when the technology gap between the leader and the follower is relatively small.[1] One may expect that full patent protection may be suboptimal in a world of step-by-step competition; by stochastically or deterministically allowing the follower to use the innovations of the technological leader, the likelihood of relatively small gap between leaders and followers, and thus the amount of R&D, may be raised.[2] Based on this intuition, one may further conjecture that state-dependent IPR policy, when feasible, should provide *less* protection to firms that are technologically more advanced relative to their competitors.

There are two problems with this intuition, however. First, it is derived from models with uniform IPR policy, where relaxation of patent protection always discourages R&D. The major contribution of our paper will be to show that state-dependent relaxation of patent protection can increase R&D. This force will lead to the *opposite* of the above conjecture and show that

---

[1] Aghion, Bloom, Blundell, Griffith and Howitt (2005) provide empirical evidence that there is greater R&D in British industries where there is a smaller technological gap between firms. See O'Donoghue, Scotchmer and Thisse (1998) for a discussion of how patent life may come to an end because of related innovations.

[2] This is conjectured, for example, in Aghion, Harris, Howitt and Vickers (2001, p. 481).

optimal IPR policy involves providing *more* protection to firms that are technologically more advanced. This is because of *trickle-down of incentives*; providing relatively low protection to firms with limited leads and greater protection to those that have greater leads not only improves the incentives of firms that are technologically advanced, but also encourages R&D by those that have limited leads because of the prospect of reaching levels of technology gaps associated with greater protection. Second, this conjecture is based on models that do not allow licensing of the leading-edge technology. Introducing licensing changes the trade-offs underlying the above intuition and the implied form of the optimal IPR policy.[3]

To investigate these issues systematically, we construct a general equilibrium model with step-by-step innovation, potential licensing of patents and state-dependent IPR policy. In our model economy, each firm can climb the technology ladder via three different methods: (i) by "catch-up R&D," that is, R&D investments applied to a variant of the technology of the leader; (ii) by "frontier R&D," that is, building on the patented innovations of the technological leader for a pre-specified *license fee*; and (iii) as a result of the expiration of the patent of the technological leader.

The presence of various different forms of technological progress in this model allows for a range of different policy regimes. The first is *full patent protection with no licensing*, which corresponds to the environment assumed in existing growth models (e.g., Aghion, Harris, Howitt and Vickers, 2001) and provides full (indefinite) patent protection to technological leaders, but does not allow any licensing agreements (it sets the license fees to infinity). The second is *full patent protection with (compulsory) licensing*. This regime allows technological followers to build on the leading-edge technology in return for a pre-specified license fee. Licensing is "compulsory" in this regime in the sense that the patent holder does not have the right to refuse to license its innovation to a follower that is willing to pay the pre-specified license fee. There is "full patent protection" in the sense that patents never expire and the license fee is equal to the gain in net present value accruing to the follower because of its use of the leading-edge technology.[4] The third regime is *uniform imperfect patent protection*, which deviates from the

---

[3] See Scotchmer (2005) for the importance of incorporating these types of licensing agreements into models of innovation.

[4] Compulsory license fees may be based on the damage that the use of the technology causes to the technological leader (because of loss of profits) or on the gain to followers from the use of superior technology. In practice,

11

previous two benchmarks by allowing either expiration of patents and/or license fees that are less than the full benefit to the follower. The adjective uniform indicates that in this policy environment all industries are treated identically regardless of the technology gap between the leader and the follower. The final and most interesting policy regime is *state-dependent imperfect patent protection*, which deviates from full patent protection as a function of the technology gap between the leader and the follower in the industry (i.e., it allows technologically more advanced firms to receive a different amount of IPR protection). Each of these policy regimes captures a different conceptualization of IPR policy and is interesting in its own right (and naturally, the last regime is general enough to nest the other three).

We first prove the existence of a stationary (steady-state) equilibrium under any of these policy regimes and characterize a number of features of the equilibrium analytically. For example, we prove that with uniform IPR policy, R&D investments decline when the gap between the leader and the follower increases.

We then turn to a quantitative investigation of welfare-maximizing ("optimal") IPR policy. We provide a simple calibration of our baseline model and then derive the optimal IPR given this economy. This calibration exercise only requires the choice of two parameters and the functional form for the R&D production function. Despite its simplicity and parsimony, the model generates reasonable numbers for the allocation of the workforce between production and research and the magnitude of profits in GDP. Our quantitative investigation leads to two major results:

1. Allowing for (compulsory) licensing of patents increases the equilibrium growth and welfare of the economy significantly. Intuitively, without such licensing, a large part of the R&D effort goes to duplication, and followers' R&D does not directly contribute to growth. Licensing implies that R&D by all firms—not just the leaders—contributes to growth and also increases the R&D incentives of followers. In our benchmark parameterization, allowing for licensing increases the steady-state equilibrium growth rate of the economy from 1.86% to 2.58% per

---

licensing fees or patent infringement fees reflect both the benefits to the firm using the knowledge and the damage to the original inventor (see, e.g., Scotchmer, 2005). In our analysis, we allow license fees to be set at any level, thus incorporating both possibilities. The analysis can also be extended to allow for bilateral licensing arrangements between the leader and the follower in the industry, for example at some license fee that results from a bargain between them. We will discuss this possibility in subsection 1.3.3 and argue that compulsory licensing fees typically improve welfare and growth relative to bilateral agreements.

annum and also has a significant effect on steady-state welfare.

2. More importantly, we show that welfare-maximizing IPR policy is state dependent and provides *greater protection to firms that are technologically more advanced* (relative to technological leaders that only have a small lead over their followers). In particular, because of the disincentive effect of relaxing IPR protection on R&D, uniform IPR policy (either by manipulating license fees or the duration of patents) has a minimal effect on growth and welfare. In contrast, state-dependent IPR policy can significantly increase innovation, growth and welfare. For example, in our baseline parameterization, optimal state-dependent IPR policy increases the growth rate to 2.96% relative to the growth rate of 2.63% under (optimal) uniform IPR policy.

The reason why optimal IPR policy provides greater protection to technological leaders that are further ahead than their rivals is the *trickle-down effect*. When a particular state for the technological leader (say being n^* steps ahead of the follower) is very profitable, this increases the incentives to perform R&D not only for leaders that are $n^* - 1$ steps ahead, but for *all* leaders with a lead of size $n \leq n^* - 1$. The trickle-down effect makes state-dependent IPR, with greater protection for firms that are technologically more advanced than their rivals, preferable to uniform IPR. Another implication of the trickle-down effect is also worth noting. As is well known, uniform relaxation of IPR protection *reduces* R&D incentives (because innovation is rewarded less). However, because of the trickle-down effect, state-dependent relaxation of IPR may *increase* (average) R&D investments—because increasing protection at technology gap n^* and reducing it at $n^* - k$ creates a big boost to the R&D of firms with technological lead of $n^* - k$ steps. We will show that for plausible parameter values the amount of R&D is greater under imperfect state-dependent IPR protection than under full IPR protection, and this will be the main reason why state-dependent IPR can have a significant positive effect on economic growth.

Our work is a contribution both to the IPR protection and the endogenous growth literatures. Previous work in industrial organization and in growth theory emphasizes that ex-post monopoly rents and thus patents are central for generating the ex-ante investments in R&D and technological progress, even though monopoly power also creates distortions (e.g., Arrow,

13

1962, Reinganum, 1981, Tirole, 1988, Romer, 1990, Grossman and Helpman, 1991, Aghion and Howitt, 1992, Green and Scotchmer, 1995, Scotchmer, 1999, Gallini and Scotchmer, 2002, O'Donoghue and Zweimuller, 2004).[5] Much of the literature discusses the trade-off between these two forces to determine the optimal length and breadth of patents. For example, Klemperer (1990) and Gilbert and Shapiro (1990) show that optimal patents should have a long duration in order to provide inducement to R&D, but a narrow breadth so as to limit monopoly distortions. A number of other papers, for example, Gallini (1992) and Gallini and Scotchmer (2002), reach opposite conclusions.

Another branch of the literature, including the seminal paper by Scotchmer (1999) and the recent interesting papers by Llobet, Hopenhayn and Mitchell (2006) and Hopenhayn and Mitchell (2001), adopts a mechanism design approach to the determination of the optimal patent and intellectual property rights protection system. For example, Scotchmer (1999) derives the patent renewal system as an optimal mechanism in an environment where the cost and value of different projects are unobserved and the main problem is to decide which projects should go ahead. Llobet, Hopenhayn and Mitchell (2006) consider optimal patent policy in the context of a model of sequential innovation with heterogeneous quality and private information. They show that allowing for a choice from a menu of patents will be optimal in this context. To the best of our knowledge, no other paper in the literature has considered state-dependent IPR policy or developed the general equilibrium framework for IPR policy analysis. As a first attempt, we only look at state-dependent patent length and license fees (though similar ideas can be applied to an investigation of the gains from making the breadth of patent awards state-dependent).

Our work is most closely related to and extends the results of Aghion, Harris and Vickers (1997) and Aghion, Harris, Howitt and Vickers (2001) on endogenous growth with step-by-step innovation.[6] Although our model builds on these papers, it also differs from them in a number of significant ways. First, we allow licensing agreements whereby followers can pay a pre-specified license fee for building on the leading-edge technology developed by other firms. We show that such licensing has significant effects on growth and welfare. Second, our economy

---

[5] Boldrin and Levine (2001, 2004) or Quah (2003) argue that patent systems are not necessary for innovation.
[6] Segal and Whinston (2005) analyze the impact of anti-trust policy on economic growth in a related model of step-by-step innovation.

incorporates a general IPR policy that can be state dependent. Third, in our economy there is a general equilibrium interaction between production and R&D, since they both compete for scarce labor.[7] Finally, we provide a number of analytical results for the general model (with or without IPR policy), while previous literature has focused on the special cases where innovations are either "drastic" (so that the leader never undertakes R&D) or very small, and has not provided existence or general characterization results for steady-state equilibria in this class of economies.

Lastly, our results are also related to the literature on tournaments and races, for example, Fudenberg, Gilbert, Stiglitz and Tirole (1983), Harris and Vickers (1985, 1987), Choi (1991), Budd, Harris and Vickers (1993), Taylor (1995), Fullerton and McAfee (1999), Baye and Hoppe (2003), and Moscarini and Squintani (2004). This literature considers the impact of endogenous or exogenous prizes on effort in tournaments, races or R&D contests. In terms of this literature, state-dependent IPR policy can be thought of as "state-dependent handicapping" of different players (where the state variable is the gap between the two players in a dynamic tournament). To the best of our knowledge, these types of schemes have not been considered in this literature.

The rest of the chapter is organized as follows. Section 1.2 presents the basic environment. Section 1.3 proves the existence of a steady-state equilibrium and characterizes some of its key properties under both uniform and state-dependent IPR policy. In this section, we also briefly discuss how bargaining over license fees can be incorporated into our framework and why compulsory licensing fees would have a useful role even in the presence of bilateral bargaining between technology leaders and followers. Section 1.4 quantitatively evaluates the implications of various different types of IPR policy regimes on welfare and characterizes the welfare-maximizing state-dependent IPR policies. Section 1.5 concludes, while the Appendix contains the proofs of all the results stated in the text.

---

[7]This general equilibrium aspect is introduced to be able to close the model economy without unrealistic assumptions and makes our economy more comparable to other growth models (Aghion, Harris, Howit and Vickers, 2001, assume a perfectly elastic supply of labor). We show that the presence of general equilibrium interactions does not significantly complicate the analysis and it is still possible to characterize the steady-state equilibrium.

1.2 Model

We now describe the basic environment. The characterization of the equilibrium under the different policy regimes is presented in the next section.

1.2.1 Preferences and Technology

Consider the following continuous time economy with a unique final good. The economy is populated by a continuum of 1 individuals, each with 1 unit of labor endowment, which they supply inelastically. Preferences at time t are given by

$$\mathbb{E}_t \int_t^\infty \exp\left(-\rho\left(s-t\right)\right) \log C\left(s\right) ds, \tag{1.1}$$

where $\mathbb{E}_t$ denotes expectations at time t, $\rho > 0$ is the discount rate and $C\left(t\right)$ is consumption at date t. The logarithmic preferences in (1.1) facilitate the analysis, since they imply a simple relationship between the interest rate, growth rate and the discount rate (see (1.2) below).

Let $Y\left(t\right)$ be the total production of the final good at time t. We assume that the economy is closed and the final good is used only for consumption (i.e., there is no investment), so that $C\left(t\right) = Y\left(t\right)$. The standard Euler equation from (1.1) then implies that

$$g\left(t\right) \equiv \frac{\dot{C}\left(t\right)}{C\left(t\right)} = \frac{\dot{Y}\left(t\right)}{Y\left(t\right)} = r\left(t\right) - \rho, \tag{1.2}$$

where this equation defines $g\left(t\right)$ as the growth rate of consumption and thus output, and $r\left(t\right)$ is the interest rate at date t.

The final good Y is produced using a continuum 1 of intermediate goods according to the Cobb-Douglas production function

$$\ln Y\left(t\right) = \int_0^1 \ln y\left(j,t\right) dj, \tag{1.3}$$

where $y\left(j,t\right)$ is the output of jth intermediate at time t. Throughout, we take the price of the final good as the numeraire and denote the price of intermediate j at time t by $p\left(j,t\right)$. We also assume that there is free entry into the final good production sector. These assumptions, together with the Cobb-Douglas production function (3.3.2), imply that the final good sector

16

has the following demand for intermediates

$$y(j,t) = \frac{Y(t)}{p(j,t)}, \quad \forall j \in [0,1].$$ (1.4)

Intermediate $j \in [0,1]$ comes in two different *varieties*, each produced by one of two infinitely-lived firms. We assume that these two varieties are perfect substitutes and these firms compete a la Bertrand.[8] Firm $i = 1$ or 2 in industry j has the following technology

$$y(j,t) = q_i(j,t)\, l_i(j,t)$$ (1.5)

where $l_i(j,t)$ is the employment level of the firm and $q_i(j,t)$ is its level of technology at time t. Each consumer in the economy holds a balanced portfolio of the shares of all firms. Consequently, the objective function of each firm is to maximize expected profits.

The production function for intermediate goods, (1.5), implies that the marginal cost of producing intermediate j for firm i at time t is

$$MC_i(j,t) = \frac{w(t)}{q_i(j,t)}$$ (1.6)

where $w(t)$ is the wage rate in the economy at time t.

When this causes no confusion, we denote the *technological leader* in each industry by i and the follower by $-i$, so that we have:

$$q_i(j,t) \geq q_{-i}(j,t).$$

Bertrand competition between the two firms implies that all intermediates will be supplied by

---

[8] A more general case would involve these two varieties being imperfect substitutes, for example, with the output of intermediate j produced as

$$y(j,t) = \left[\varphi y_1(j,t)^{\frac{\sigma-1}{\sigma}} + (1-\varphi) y_2(j,t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}},$$

with $\sigma > 1$. The model analyzed in the text corresponds to the limiting case where $\sigma \to \infty$. Our results can be easily extended to this more general case with any $\sigma > 1$, but at the cost of additional notation. We therefore prefer to focus on the case where the two varieties are perfect substitutes. It is nonetheless useful to bear this formulation with imperfect substitutes in mind, since it facilitates the interpretation of "distinct" innovations by the two firms (when the follower engages in "catch-up" R&D).

the leader at the "limit" price:[9]

$$p_i(j, t) = \frac{w(t)}{q_{-i}(j, t)}.$$ (1.7)

Equation (1.4) then implies the following demand for intermediates:

$$y(j, t) = \frac{q_{-i}(j, t)}{w(t)} Y(t).$$ (1.8)

1.2.2 Technology, R&D and IPR Policy

R&D by the leader or the follower stochastically leads to innovation. We assume that when the leader innovates, its technology improves by a factor $\lambda > 1$.

The follower, on the other hand, can undertake R&D to catch up with the frontier technology or to improve over the frontier technology.[10] The first possibility is *catch-up R&D* and can be thought of R&D to discover an *alternative* way of performing the same task as the current leading-edge technology. Because this innovation applies to the follower's variant of the product (recall footnote 8) and results from its own R&D efforts, we assume that it does not constitute infringement on the patent of the leader, and the follower does not have to make any license fee payments. Therefore, if the follower chooses the first possibility, it will have to retrace the steps of the technological leader (corresponding to its own variant of the product), but in return, it will not have to pay a patent license fee. For follower firm $-i$ in industry j at time t, we denote this type of R&D by

$$a_{-i}(j, t) = 0.$$

The alternative, *frontier R&D*, involves followers building on and improving the current leading-edge technology. If this type of R&D succeeds, the follower will have improved the leading-edge technology using the patented knowledge of the technological leader, and thus will

---

[9]If the leader were to charge a higher price, then the market would be captured by the follower earning positive profits. A lower price can always be increased while making sure that all final good producers still prefer the intermediate supplied by the leader i rather than that by the follower $-i$, even if the latter were supplied at marginal cost. Since the monopoly price with the unit elastic demand curve is infinite, the leader always gains by increasing its price, making the price given in (1.7) the unique equilibrium price.

[10]A third possibility is for the follower to climb the technology ladder step-by-step, meaning that, for example, when the current leader is at some technology rung $n_{ij}(t)$ and the follower itself is at $n_{-ij}(t) < n_{ij}(t) - 1$, it must first discover technology $n_{-ij}(t) + 1$, et cetera. We have investigated this type of environment with "slow catch-up" in a previous version of the paper. Since the general results are similar, we do not discuss this variation to save space.

have to pay a license fee to the leader. The license fees may result from bargaining between the leader and the follower or they may be compulsory license fees imposed by policy. In either case, one must first characterize the equilibrium for a given sequence of license fees, which will be the main part of our analysis. We specify the license fees and how they vary below, and consistent with our main focus, we refer to them as "policy," though we will also discuss how they can be determined via bargaining.[11] This strategy is denoted by

$$a_{-i}(j, t) = 1.$$

Throughout, we allow $a_{-i}(j, t) \in [0, 1]$ for mathematical convenience, thus a should be interpreted as the probability of frontier R&D by the follower.

R&D by the leader, catch-up R&D by the follower, and frontier R&D by the follower may have different costs and success probabilities. We simplify the analysis by assuming that all three types of R&D have the same costs and the same probability of success. In particular, in all cases, we assume that innovations follow a controlled Poisson process, with the arrival rate determined by R&D investments. Each firm (in every industry) has access to the following R&D technology:

$$x_i(j, t) = F(h_i(j, t)), \tag{1.9}$$

where $x_i(j, t)$ is the flow rate of innovation at time t and $h_i(j, t)$ is the number of workers hired by firm i in industry j to work in the R&D process at t. This specification implies that within a time interval of Δt, the probability of innovation for this firm is $x_i(j, t) \Delta t + o(\Delta t)$.

We assume that F is twice continuously differentiable and satisfies $F'(\cdot) > 0$, $F''(\cdot) < 0$, $F'(0) < \infty$ and that there exists $\bar{h} \in (0, \infty)$ such that $F'(h) = 0$ for all $h \geq \bar{h}$. The assumption that $F'(0) < \infty$ implies that there is no Inada condition when $h_i(j, t) = 0$. The last assumption, on the other hand, ensures that there is an upper bound on the flow rate of innovation (which is not essential but simplifies the proofs). Recalling that the wage rate for labor is $w(t)$, the

---

[11]It should already be noted that the follower will never license the technology of the leader for production purposes, since this would lead to Bertrand competition and zero ex post profits for both parties.

cost for R&D is therefore $w(t) G(x_i(j, t))$ where

$$G(x_i(j, t)) \equiv F^{-1}(x_i(j, t)), \tag{1.10}$$

and the assumptions on F immediately imply that G is twice continuously differentiable and satisfies $G'(\cdot) > 0$, $G''(\cdot) > 0$, $G'(0) > 0$ and $\lim_{x \to \bar{x}} G'(x) = \infty$, where

$$\bar{x} \equiv F(\bar{h}) \tag{1.11}$$

is the maximal flow rate of innovation (with $\bar{h}$ defined above).

We next describe the evolution of technologies within each industry. Suppose that leader i in industry j at time t has a technology level of

$$q_i(j, t) = \lambda^{n_{ij}(t)}, \tag{1.12}$$

and that the follower $-i$'s technology at time t is

$$q_{-i}(j, t) = \lambda^{n_{-ij}(t)}, \tag{1.13}$$

where $n_{ij}(t) \geq n_{-ij}(t)$ and $n_{ij}(t)$, $n_{-ij}(t) \in \mathbb{Z}_+$ denote the technology rungs of the leader and the follower in industry j. We refer to $n_j(t) \equiv n_{ij}(t) - n_{-ij}(t)$ as the *technology gap* in industry j. If the leader undertakes an innovation within a time interval of Δt, then its technology increases to $q_i(j, t + \Delta t) = \lambda^{n_{ij}t + 1}$ and the technology gap rises to $n_j(t + \Delta t) = n_j(t) + 1$ (the probability of two or more innovations within the interval Δt will be $o(\Delta t)$, where $o(\Delta t)$ represents terms that satisfy $\lim_{\Delta t \to 0} o(\Delta t) / \Delta t$).

When the follower is successful in catch-up R&D (i.e., $a_{-i}(j, t) = 0$) within the interval Δt, then its technology improves to

$$q_{-i}(j, t + \Delta t) = \lambda^{n_{ij}t},$$

and the technology gap variable becomes $n_{jt + \Delta t} = 0$. In contrast, if the follower is successful in frontier R&D and pays the license fee (i.e., $a_{-i}(j, t) = 1$), then it surpasses the leading-edge

technology, so we have

$$q_{-i}\left(j, t + \Delta t\right) = \lambda^{n_{ijt}+1}$$

and the technology gap variable becomes $n_{jt+\Delta t} = 1$ (and from this point onwards, the labels i and $-i$ are swapped, since the previous follower now becomes the leader).

In addition to catching up with or surpassing the technology frontier with their own R&D, followers can also copy the technology frontier because IPR policy is such that some patents expire. In particular, we assume that patents expire at some policy-determined Poisson rate η, and after expiration, followers can costlessly copy the frontier technology, jumping to $q_{-i}\left(j, t + \Delta t\right) = \lambda^{n_{ijt}}.$[12]

This description makes it clear that there are two aspects to IPR policy: (i) the length of the patent (modeled as a Poisson rate of arrival of the termination of the patent); (ii) the license fees. We allow both of these to be state dependent, so they are represented by the following two functions:

$$\boldsymbol{\eta} : \mathbb{N} \to \mathbb{R}_+$$

and for all $t \geq 0$,

$$\hat{\boldsymbol{\zeta}}\left(t\right) : \mathbb{N} \to \mathbb{R}_+ \cup \left\{+\infty\right\}.$$

Here $\eta\left(n\right) \equiv \eta_n < \infty$ is the flow rate at which the patent protection is removed from a technological leader that is n steps ahead of the follower. When $\eta_n = 0$, this implies that there is full protection at technology gap n, in the sense that patent protection will never be removed. In contrast, $\eta_n \to \infty$ implies that patent protection is removed immediately once technology gap n is reached. Similarly, $\hat{\zeta}\left(n, t\right) \equiv \hat{\zeta}_n\left(t\right)$ denotes the patent fee that a follower has to pay in order to build upon the innovation of the technological leader, when the technology gap in the industry is n steps.[13] Our formulation imposes that $\boldsymbol{\eta} \equiv \left\{\eta_1, \eta_2, ...\right\}$ is time-invariant, while

---

[12] Alternative modeling assumptions on IPR policy, such as a fixed patent length of $T > 0$ from the time of innovation, are not tractable, since they lead to value functions that take the form of delayed differential equations.

[13] Throughout, we assume that ζ is a policy choice and firms cannot contract around it. An alternative approach would be to allow firms to bargain over the level of license fees. In this case, it is plausible to presume that the legally-specified infringement penalties or license fees will affect the equilibrium in the bargaining game, so the effect of policies we investigate would still be present. We do not allow bargaining between firms over the license fees in order to simplify the analysis.

$\hat{\zeta}(t) \equiv \left\{ \hat{\zeta}_1(t), \hat{\zeta}_2(t), \ldots \right\}$ is a function of time. This is natural, since in a growing economy, license fees should not remain constant. Below, we will require that $\hat{\zeta}$ grows at the same rate as aggregate output in the economy.

When $\hat{\zeta}_n(t) = 0$, there is no protection because followers can license the leading-edge technology at zero cost.[14] In contrast, when $\hat{\zeta}_n(t) = \infty$, licensing the leading-edge technology is prohibitively costly. Note however that $\hat{\zeta}_n < \infty$ does not necessarily imply that patent protection is imperfect. In particular, in what follows we interpret a situation in which the license fee is equal to the net extra gain from surpassing the leader rather than being neck-and-neck (i.e., being at a technology gap of 0) as "full protection".[15] We also refer to a policy regime as *uniform* IPR protection if both η and $\hat{\zeta}(t)$ are constant functions of n, meaning that intellectual property law treats all firms and industries identically regardless of the technology gap between the leader and the follower (i.e., $\eta^{uni} \equiv \{\eta, \eta, \ldots\}$ and $\hat{\zeta}^{uni}(t) \equiv \left\{ \hat{\zeta}(t), \hat{\zeta}(t), \ldots \right\}$). We also assume that there exists some $\bar{n} < \infty$ such that $\eta_n = \eta_{\bar{n}}$ and $\hat{\zeta}_n(t) = \hat{\zeta}_{\bar{n}}(t)$ for all $n \geq \bar{n}$.

Given this specification, we can now write the law of motion of the technology gap in

---

[14]Throughout, we interpret $\zeta_n(t) = 0$ as $\zeta_n(t) = \varepsilon$ with $\varepsilon \downarrow 0$, so that followers continue not to license the new technology without innovation (recall the comment in footnote 11).

[15]In other words, we interpret "full protection" to correspond to a situation in which $\zeta_n(t) \geq V_1(t) - V_0(t)$, where V_1 refers to the net present value of a firm that is one step ahead of its rival and V_0 is the value of a firm that is neck-and-neck with its rival. Alternatively, full protection could be interpreted as corresponding to the case in which the follower pays a license fee equal to the loss of profits that it causes for the technology leader (see Scotchmer, 2005). In our model, this would correspond to $\zeta_n(t) = V_0(t) - V_{-1}(t)$, where V_{-1} is the net present value of a firm that is one step behind the technology leader. In all equilibria we compute below, we find that the second amount is significantly less than the first, thus our notion of full protection licensing fee is large enough to cover both possibilities. In any case, what value of ζ is designated as "full protection" does not have any bearing on our formal analysis, since we characterize the equilibrium for any ζ and then find the welfare-maximizing policy sequence.

industry j as follows:

$$
n_j(t + \Delta t) = \begin{cases} n_j(t) + 1 & \text{with probability} \quad x_i(j,t)\,\Delta t + o(\Delta t) \\ \\ 0 & \text{with probability} \quad \left((1 - a_{-i}(j,t))\,x_{-i}(j,t) + \eta_{n_j(t)}\right)\Delta t + o(\Delta t) \\ \\ 1 & \text{with probability} \quad a_{-i}(j,t)\,x_{-i}(j,t)\,\Delta t + o(\Delta t) \\ \\ n_j(t) & \text{with probability} \quad 1 - \left(x_i(j,t) + x_{-i}(j,t) + \eta_{n_j(t)}\right)\Delta t - o(\Delta t)) \end{cases}
$$
$$\tag{1.14}$$

Here $o(\Delta t)$ again represents second-order terms, in particular, the probabilities of more than one innovations within an interval of length Δt. The terms $x_i(j,t)$ and $x_{-i}(j,t)$ are the flow rates of innovation by the leader and the follower; $a_{-i}(j,t) \in [0,1]$ denotes whether the follower is trying to catch up with a leader or surpass it; and $\eta_{n_j(t)}$ is the flow rate at which the follower is allowed to copy the technology of a leader that is $n_j(t)$ steps ahead. Intuitively, the technology gap in industry j increases from $n_j(t)$ to $n_j(t)+1$ if the leader is successful. When $a_{-i}(j,t) = 1$, the technology gap in industry j becomes 1 if the follower is successful (flow rate $x_{-i}(j,t)$). Finally, the firms become "neck-and-neck" when the follower comes up with an alternative technology to that of the leader (flow rate $x_{-i}(j,t)$) without using the license ($a_{-i}(j,t) = 0$) or the patent expires at the flow rate η_{n_j}.

1.2.3 Profits

We next write the instantaneous "operating" profits for the leader (i.e., the profits exclusive of R&D expenditures and license fees). Profits of leader i in industry j at time t are

$$
\begin{aligned}
\Pi_i(j,t) &= [p_i(j,t) - MC_i(j,t)]\,y_i(j,t) \\
&= \left(\frac{w(t)}{q_{-i}(j,t)} - \frac{w(t)}{q_i(j,t)}\right)\frac{Y(t)}{p_i(j,t)} \\
&= \left(1 - \lambda^{-n_j(t)}\right)Y(t)
\end{aligned}
\tag{1.15}
$$

23

where $n_j(t) \equiv n_{ij}(t) - n_{-ij}(t)$ is the technology gap in industry j at time t. The first line simply uses the definition of operating profits as price minus marginal cost times quantity sold. The second line uses the fact that the equilibrium limit price of firm i is $p_i(j,t) = w(t)/q_{-i}(j,t)$ as given by (1.7), and the final equality uses the definitions of $q_i(j,t)$ and $q_{-i}(j,t)$ from (1.12) and (1.13). The expression in (1.15) also implies that there will be zero profits in neck-and-neck industries, i.e., in those with $n_j(t) = 0$. Also clearly, followers always make zero profits, since they have no sales.

The Cobb-Douglas aggregate production function in (3.3.2) is responsible for the form of the profits (1.15), since it implies that profits only depend on the technology gap of the industry and aggregate output. This will simplify the analysis below by making the technology gap in each industry the only industry-specific payoff-relevant state variable.

The objective function of each firm is to maximize the net present discounted value of "net profits" (operating profits minus R&D expenditures and plus or minus patent fees). In doing this, each firm will take the sequence of interest rates, $[r(t)]_{t \geq 0}$, the sequence of aggregate output levels, $[Y(t)]_{t \geq 0}$, the sequence of wages, $[w(t)]_{t \geq 0}$, the R&D decisions of all other firms and policies as given.

1.2.4 Equilibrium

Let $\mu(t) \equiv \{\mu_n(t)\}_{n=0}^{\infty}$ denote the distribution of industries over different technology gaps, with $\sum_{n=0}^{\infty} \mu_n(t) = 1$. For example, $\mu_0(t)$ denotes the fraction of industries in which the firms are neck-and-neck at time t. Throughout, we focus on Markov Perfect Equilibria (MPE), where strategies are only functions of the payoff-relevant state variables.[16] This allows us to drop the dependence on industry j, thus we refer to R&D decisions by x_n for the technological leader that is n steps ahead and by a_{-n} and x_{-n} for a follower that is n steps behind. Let us denote the list of decisions by the leader and the follower with technology gap n at time t by $\xi_n(t) \equiv \langle x_n(t), p_i(j,t), y_i(j,t) \rangle$ and $\xi_{-n}(t) \equiv \langle a_{-n}(t), x_{-n}(t) \rangle$.[17] Throughout, $\boldsymbol{\xi}$ will indicate

---

[16] MPE is a natural equilibrium concept in this context, since it does not allow for implicit collusive agreements between the follower and the leader. While such collusive agreements may be likely when there are only two firms in the industry, in most industries there are many more firms and also many potential entrants, making collusion more difficult. Throughout, we assume that there are only two firms to keep the model tractable.

[17] The price and output decisions, $p_i(j,t)$ and $y_i(j,t)$, depend not only on the technology gap, aggregate output and the wage rate, but also on the exact technology rung of the leader, $n_{ij}(t)$. With a slight abuse of notation,

the whole sequence of decisions at every state, so that $\boldsymbol{\xi}(t) \equiv \{\xi_n(t)\}_{n=-\infty}^{\infty}$. We define an allocation as follows:

Definition 1.1 *(Allocation)* *Let* $< \eta, \ [\hat{\boldsymbol{\zeta}}(t)]_{t \geq 0} >$ *be the IPR policy sequences. Then an allocation is a sequence of decisions for a leader that is* $n = 0, 1, 2, \ldots$ *step ahead,* $[\xi_n(t)]_{t \geq 0}$, *a sequence of R&D decisions for a follower that is* $n = 1, 2, \ldots$ *step behind,* $[\xi_{-n}(t)]_{t \geq 0}$, *a sequence of wage rates* $[w(t)]_{t \geq 0}$, *and a sequence of industry distributions over technology gaps* $[\boldsymbol{\mu}(t)]_{t \geq 0}$.

For given IPR sequences η and $[\hat{\boldsymbol{\zeta}}(t)]_{t \geq 0}$, MPE strategies, which are only functions of the payoff-relevant state variables, can be represented as follows

$$\mathbf{x} \ : \ \mathbb{Z} \times \mathbb{R}_+^2 \times [0,1]^\infty \to \mathbb{R}_+,$$

$$a \ : \ \mathbb{Z}_- \backslash \{0\} \times \mathbb{R}_+^2 \times [0,1]^\infty \to [0,1].$$

The first mapping represents the R&D decision of a firm (both when it is the follower and when it is the leader in an industry) as a function of the technology gap, $n \in \mathbb{Z}$, the aggregate level of output and the wage, $(Y, w) \in \mathbb{R}_+^2$, and R&D decision of the other firm in the industry, $\tilde{\mathbf{x}} \in [0,1]^\infty$. The second function represents the follower's decision of whether to direct its R&D to catching up with or surpassing the leading-edge technology (or more precisely, it represents the probability with which the follower will choose to undertake R&D to surpass the leading edge technology). Consequently, we have the following definition of equilibrium:

Definition 1.2 *(Equilibrium)* *Given an IPR policy sequence* $< \eta, \ [\hat{\boldsymbol{\zeta}}(t)]_{t \geq 0} >$, *a Markov Perfect Equilibrium is given by a sequence* $[\boldsymbol{\xi}^*(t), w^*(t), Y^*(t)]_{t \geq 0}$ *such that (i)* $[p_i^*(j,t)]_{t \geq 0}$ *and* $[y_i^*(j,t)]_{t \geq 0}$ *implied by* $[\boldsymbol{\xi}^*(t)]_{t \geq 0}$ *satisfy (1.7) and (1.8); (ii) R&D policies* $[a^*(t), \mathbf{x}^*(t)]_{t \geq 0}$ *are best responses to themselves, i.e.,* $[a^*(t), \mathbf{x}^*(t)]_{t \geq 0}$ *maximizes the expected profits of firms taking aggregate output* $[Y^*(t)]_{t \geq 0}$, *wages* $[w^*(t)]_{t \geq 0}$, *government policy* $< \eta, \ [\hat{\boldsymbol{\zeta}}(t)]_{t \geq 0} >$ *and the R&D policies of other firms* $[a^*(t), \mathbf{x}^*(t)]_{t \geq 0}$ *as given; (iii) aggregate output* $[Y^*(t)]_{t \geq 0}$ *is given by (3.3.2); and (iv) the labor market clears at all times given the wage sequence* $[w^*(t)]_{t \geq 0}$.

---

throughout we suppress this dependence, since their product $p_i(j,t) y_i(j,t)$ and the resulting profits for the firm, (1.15), are independent of $n_{ij}(t)$, and consequently, only the technology gap, $n_j(t)$, matters for profits, R&D, aggregate output and economic growth.

1.2.5 The Labor Market

Since only the technological leader produces, labor demand in industry j with technology gap $n_j(t) = n$ can be expressed as

$$l_n(t) = \frac{\lambda^{-n} Y(t)}{w(t)} \quad \text{for} \quad n \in \mathbb{Z}_+. \tag{1.16}$$

In addition, there is demand for labor coming for R&D from both followers and leaders in all industries. Using (1.9) and the definition of the G function, we can express industry demands for R&D labor as

$$h_n(t) = G(x_n(t)) + G(x_{-n}(t)) \quad \text{for} \quad n \in \mathbb{Z}_+ \;, \tag{1.17}$$

where $G(x_n(t))$ and $G(x_{-n}(t))$ refer to the demand of the leader and the follower in an industry with a technology gap of n. Note that in this expression, $x_{-n}(t)$ refers to the R&D effort of a follower that is n steps behind (conditional on its optimal choice of $a_{-n}(t) \in [0,1]$).

The labor market clearing condition can then be expressed as:

$$1 \geq \sum_{n=0}^{\infty} \mu_n(t) \left[\frac{1}{\omega(t)\lambda^n} + G(x_n(t)) + G(x_{-n}(t)) \right], \tag{1.18}$$

and $\omega(t) \geq 0$, with complementary slackness, where

$$\omega(t) \equiv \frac{w(t)}{Y(t)} \tag{1.19}$$

is the labor share at time t. The labor market clearing condition, (1.18), uses the fact that total supply is equal to 1, and demand cannot exceed this amount. If demand falls short of 1, then the wage rate, $w(t)$, and thus the labor share, $\omega(t)$, have to be equal to zero (though this will never be the case in equilibrium). The right-hand side of (1.18) consists of the demand for production (the terms with ω in the denominator), the demand for R&D workers from the neck-and-neck industries ($2G(x_0(t))$ when $n=0$) and the demand for R&D workers coming from leaders and followers in other industries ($G(x_n(t)) + G(x_{-n}(t))$ when $n > 0$).

Defining the index of aggregate quality in this economy by the aggregate of the qualities of

the leaders in the different industries, i.e.,

$$\ln Q(t) \equiv \int_0^1 \ln q_i(j, t)\, dj, \tag{1.20}$$

the equilibrium wage can be written as:[18]

$$w(t) = Q(t)\, \lambda^{-\sum_{n=0}^{\infty} n\mu_n(t)}. \tag{1.21}$$

1.2.6 Steady State and the Value Functions

Let us now focus on steady-state (Markov Perfect) equilibria, where the distribution of industries $\mu(t) \equiv \{\mu_n(t)\}_{n=0}^{\infty}$ is stationary, $\omega(t)$ defined in (1.19) and g, the growth rate of the economy, are constant over time. We will establish the existence of such an equilibrium and characterize a number of its properties. If the economy is in steady state at time $t = 0$, then by definition, we have $Y^*(t) = Y_0 e^{g^* t}$ and $w^*(t) = w_0 e^{g^* t}$, where g^* is the steady-state growth rate. These two equations also imply that $\omega(t) = \omega^*$ for all $t \geq 0$. Throughout, we assume that the parameters are such that the steady-state growth rate g^* is positive but not large enough to violate the transversality conditions. This implies that net present values of each firm at all points in time will be finite. This enables us to write the maximization problem of a leader that is $n > 0$ steps ahead recursively.

First note that given an optimal policy $\hat{x}$ for a firm, the net present discounted value of a leader that is n steps ahead at time t can be written as:

$$V_n(t) = \mathbb{E}_t \int_t^\infty \exp\left(-r(s - t)\right) \left[\Pi(s) + Z(s) - w(s)\, G(\hat{x}(s))\right] ds$$

where $\Pi(s)$ is the operating profit at time $s \geq t$, $Z(s)$ is the patent license fees received (or paid) by a firm which is the leader and $w(s)\, G(\hat{x}(s))$ denotes the R&D expenditure at time $s \geq t$. All variables are stochastic and depend on the evolution of the technology gap within the industry.

---

[18] Note that $\ln Y(t) = \int_0^1 \ln q_i(j, t)\, l(j, t)\, dj = \int_0^1 \left[\ln q_i(j, t) + \ln \frac{Y(t)}{w(t)} \lambda^{-n_j}\right] dj$, where the second equality uses (1.16). Thus we have $\ln Y(t) = \int_0^1 \left[\ln q_i(j, t) + \ln Y(t) - \ln w(t) - n_j \ln \lambda\right] dj$. Rearranging and canceling terms, and writing $\exp \int n_j \ln \lambda\, dj = \lambda^{-\sum_{n=0}^{\infty} n\mu_n(t)}$, we obtain (1.21).

Next taking as given the equilibrium R&D policy of other firms, $x^*_{-n}(t)$ and $a^*_{-n}(t)$, the equilibrium interest and wage rates, $r^*(t)$ and $w^*(t)$, and equilibrium profits $\{\Pi^*_n(t)\}_{n=1}^\infty$ (as a function of equilibrium aggregate output), this value can be written as (see the Appendix for the derivation of this equation):[19]

$$r^*(t) V_n(t) - \dot{V}_n(t) = \max_{x_n(t)} \left\{ \begin{array}{l} [\Pi^*_n(t) - w^*(t) G(x_n(t))] + x_n(t) [V_{n+1}(t) - V_n(t)] \\ + \left(\left(1 - a^*_{-n}(t)\right) x^*_{-n}(t) + \eta_n\right) [V_0(t) - V_n(t)] \\ + \left(a^*_{-n}(t) x^*_{-n}(t) + \eta_n\right) \left[V_{-1}(t) - V_n(t) + \hat{\zeta}_n\right] \end{array} \right\}, \quad (1.22)$$

where $\dot{V}_n(t)$ denotes the derivative of $V_n(t)$ with respect to time. The first term is current profits minus R&D costs, while the second term captures the fact that the firm will undertake an innovation at the flow rate $x_n(t)$ and increase its technology lead by one step. The remaining terms incorporate changes in value due to catch-up by the follower (flow rate $\left(1 - a^*_{-n}(t)\right) x^*_{-n}(t) + \eta_n$ in the second line) and due to the follower leapfrogging the leader (flow rate $a^*_{-n}(t) x^*_{-n}(t)$ in the third line). In this last case, the follower will make a payment of $\hat{\zeta}_n$ to the leader for the license.

In steady state, the net present value of a firm that is n steps ahead, $V_n(t)$, will also grow at a constant rate g^* for all $n \in \mathbb{Z}_+$. Let us then define the normalized values as

$$v_n(t) \equiv \frac{V_n(t)}{Y(t)} \quad (1.23)$$

for all $n \in \mathbb{Z}$, which will be independent of time in steady state, i.e., $v_n(t) = v_n$. Similarly, in what follows we assume that license fees are also scaled up by GDP, so that

$$\zeta_n \equiv \frac{\hat{\zeta}_n(t)}{Y(t)},$$

which will ensure the existence of a (stationary) steady-state equilibrium.

Using (1.23) and the fact that from (1.2), $r(t) = g(t) + \rho$, the recursive form of the steady-

---

[19]Clearly, this value function could be written for any arbitrary sequence of R&D policies of other firms. We set the R&D policies of other firms to their equilibrium values, $x^*_{-n}(t)$ and $a^*_{-n}(t)$, to reduce notation in the main body of the paper.

state value function (1.22) can be written as:

$$\rho v_n = \max_{x_n} \left\{ \begin{array}{c} \left(1 - \lambda^{-n}\right) - \omega^* G\left(x_n\right) + x_n \left[v_{n+1} - v_n\right] \\ + \left[\left(1 - a_{-n}^*\right) x_{-n}^* + \eta_n\right] \left[v_0 - v_n\right] + a_{-n}^* x_{-n}^* \left[v_{-1} - v_n + \zeta_n\right] \end{array} \right\} \text{ for } n \in \mathbb{N},$$

(1.24)

where x_{-n}^* is the equilibrium value of R&D by a follower that is n steps behind, and ω^* is the steady-state labor share (while x_n is now explicitly chosen to maximize v_n).

Similarly the value for neck-and-neck firms is

$$\rho v_0 = \max_{x_0} \left\{ -\omega^* G\left(x_0\right) + x_0 \left[v_1 - v_0\right] + x_0^* \left[v_{-1} - v_0\right] \right\},$$

(1.25)

while the values for followers are given by

$$\rho v_{-n} = \max_{x_{-n}, a_{-n}} \left\{ \begin{array}{c} -\omega^* G\left(x_{-n}\right) + \left[\left(1 - a_{-n}\right) x_{-n} + \eta_n\right] \left[v_0 - v_{-n}\right] \\ + a_{-n} x_{-n} \left[v_1 - v_{-n} - \zeta_n\right] + x_n^* \left[v_{-n-1} - v_{-n}\right] \end{array} \right\} \text{ for } n \in \mathbb{N},$$

(1.26)

which takes into account that if the follower decides to build upon the leading-edge technology, when it innovates it will become the new leader but will have to pay the patent fee ζ_n.

For neck-and-neck firms and followers, there are no instantaneous profits, which is reflected in (1.25) and (1.26). In the former case this is because neck-and-neck firms sell at marginal cost, and in the latter case, this is because followers have no sales. These normalized value functions emphasize that, because of growth, the effective discount rate is $r\left(t\right) - g\left(t\right) = \rho$ rather than $r\left(t\right)$.

The maximization problems in (1.24)-(1.25) immediately imply that any steady-state equilibrium R&D policies, $\langle a^*, x^* \rangle$, must satisfy:

$$a_{-n}^* \left\{ \begin{array}{ll} = 1 & \text{if } v_1 - \zeta_n > v_0 \\ \in [0, 1] & \text{if } v_1 - \zeta_n = v_0 \\ = 0 & \text{if } v_1 - \zeta_n < v_0 \end{array} \right.$$

(1.27)

and

$$x_n^* = \max\left\{ G'^{-1}\left(\frac{[v_{n+1} - v_n]}{\omega^*}\right), 0 \right\} \tag{1.28}$$

$$x_{-n}^* = \max\left\{ G'^{-1}\left(\frac{(1 - a_{-n}^*)[v_0 - v_{-n}] + a_{-n}^*[v_1 - v_{-n} - \zeta_n]}{\omega^*}\right), 0 \right\} \tag{1.29}$$

$$x_0^* = \max\left\{ G'^{-1}\left(\frac{[v_1 - v_0]}{\omega^*}\right), 0 \right\}, \tag{1.30}$$

where the normalized value functions, the vs, are evaluated at the equilibrium, and $G'^{-1}(\cdot)$ is the inverse of the derivative of the G function. Since G is twice continuously differentiable and strictly concave, G'^{-1} is continuously differentiable and strictly increasing. These equations therefore imply that innovation rates, the x_n^*s, will increase whenever the incremental value of moving to the next step is greater and when the cost of R&D, as measured by the normalized wage rate, ω^*, is less. Note also that since $G'(0) > 0$, these R&D levels can be equal to zero, which is taken care of by the max operator.

The response of innovation rates, x_n^*, to the increments in values, $v_{n+1} - v_n$, is the key economic force in this model. For example, a policy that reduces the patent protection of leaders that are $n+1$ steps ahead (by increasing η_{n+1} or reducing ζ_{n+1}) will make being $n+1$ steps ahead less profitable, thus reduce $v_{n+1} - v_n$ and x_n^*. This corresponds to the standard *disincentive effect* of relaxing IPR policy. In contrast to existing models, however, here relaxing IPR policy can also create a *positive incentive effect*. This novel incentive effect has two components. First, as equation (1.29) shows, weaker patent protection in the form of lower license fees (lower ζ) may encourage further frontier R&D by the followers, directly contributing to aggregate growth. Second and perhaps somewhat more paradoxically, lower protection for technological leaders that are $n + 1$ steps ahead will tend to reduce v_{n+1}, thus increasing $v_{n+2} - v_{n+1}$ and x_{n+1}^*. We will see that this latter effect plays an important role in the form of optimal state-dependent IPR policy. In addition to the incentive effects, relaxing IPR protection may also create a beneficial *composition effect*; this is because, typically, $\{v_{n+1} - v_n\}_{n=0}^{\infty}$ is a decreasing sequence, which implies that x_{n-1}^* is higher than x_n^* for $n \geq 1$ (see, e.g., Proposition 1.2). Weaker patent protection (in the form of shorter patent lengths) will shift more industries into the neck-and-neck state and potentially increase the equilibrium level of R&D in the economy. Finally, weaker

patent protection also creates a beneficial "level effect" by influencing equilibrium markups and prices (as shown in equation (1.7) above) and by reallocating some of the workers engaged in "duplicative" R&D to production. This level effect will also feature in our welfare computations. The optimal level and structure of IPR policy in this economy will be determined by the interplay of these various forces.

Given the equilibrium R&D decisions $\langle a^*, \mathbf{x}^* \rangle$, the steady-state distribution of industries across states $\boldsymbol{\mu}^*$ has to satisfy the following accounting identities:

$$\left(x^*_{n+1} + x^*_{-n-1} + \eta_{n+1} \right) \mu^*_{n+1} = x^*_n \mu^*_n \text{ for } n \in \mathbb{N}, \tag{1.31}$$

$$\left(x^*_1 + x^*_{-1} + \eta_1 \right) \mu^*_1 = 2 x^*_0 \mu^*_0 + \sum_{n=1}^{\infty} a^*_{-n} x^*_{-n} \mu^*_n, \tag{1.32}$$

$$2 x^*_0 \mu^*_0 = \sum_{n=1}^{\infty} \left(\left(1 - a^*_{-n} \right) x^*_{-n} + \eta_n \right) \mu^*_n. \tag{1.33}$$

The first expression equates exit from state $n + 1$ (which takes the form of the leader going one more step ahead or the follower catching up for surpassing the leader) to entry into the state (which takes the form of a leader from state n making one more innovation). The second equation, (1.32), performs the same accounting for state 1, taking into account that entry into this state comes from innovation by either of the two firms that are competing neck-and-neck and also from followers that perform frontier R&D. Finally, equation (1.33) equates exit from state 0 with entry into this state, which comes from innovation by a follower in any industry with $n \geq 1$.

The labor market clearing condition in steady state can then be written as

$$1 \geq \sum_{n=0}^{\infty} \mu^*_n \left[\frac{1}{\omega^* \lambda^n} + G\left(x^*_n \right) + G\left(x^*_{-n} \right) \right] \text{ and } \omega^* \geq 0, \tag{1.34}$$

with complementary slackness.

The next proposition characterizes the steady-state growth rate. As with all the other results in the chapter, the proof of this proposition is provided in the Appendix.

Proposition 1.1 *Let the steady-state distribution of industries and R&D decisions be given by*

31

$< \boldsymbol{\mu}^*, \ \boldsymbol{a}^*, \ \mathbf{x}^* >$, *then the steady-state growth rate is*

$$g^* = \ln \lambda \left[2\mu_0^* x_0^* + \sum_{n=1}^{\infty} \mu_n^* \left(x_n^* + a_{-n}^* x_{-n}^* \right) \right].$$

(1.35)

This proposition clarifies that the steady-state growth rate of the economy is determined by three factors:

1. R&D decisions of industries at different levels of technology gap, $\mathbf{x}^* \equiv \{x_n^*\}_{n=-\infty}^{\infty}$.

2. The distribution of industries across different technology gaps, $\boldsymbol{\mu}^* \equiv \{\mu_n^*\}_{n=0}^{\infty}$.

3. Whether followers are undertaking R&D to catch up with the frontier or to surpass the frontier, $\boldsymbol{a}^* \equiv \{a_n^*\}_{n=-\infty}^{-1}$.

IPR policy affects these three margins in different directions as illustrated by the discussion above.

1.3 Existence and Characterization of Steady-State Equilibria

We now define a steady-state equilibrium in a more convenient form, which will be used to establish existence and derive some of the properties of the equilibrium.

Definition 1.3 *(Steady-State Equilibrium) Given an IPR policy $< \eta, \ \zeta >$, a steady-state equilibrium is a tuple $< \boldsymbol{\mu}^*, \ \mathbf{v}, \ \boldsymbol{a}^*, \ \mathbf{x}^*, \ \omega^*, \ g^* >$ such that the distribution of industries $\boldsymbol{\mu}^*$ satisfy (1.31), (1.32) and (1.33), the values $\mathbf{v} \equiv \{v_n\}_{n=-\infty}^{\infty}$ satisfy (1.24), (1.25) and (1.26), the R&D decisions $\boldsymbol{a}^*$ and $\mathbf{x}^*$ are given by (1.27), (1.28), (1.29) and (1.30), the steady-state labor share ω^* satisfies (1.34) and the steady-state growth rate g^* is given by (1.35).*

We next provide a characterization of the steady-state equilibrium, starting first with the case in which there is uniform IPR policy.

1.3.1 Uniform IPR Policy

Let us first focus on the case where IPR policy is uniform. This means $\eta_n = \eta < \infty$ and $\zeta_n = \zeta < \infty$ for all $n \in \mathbb{N}$ and we denote these by η^{uni} and ζ^{uni}. In this case, (1.26) implies

that the problem is identical for all followers, so that $v_{-n} = v_{-1}$ for $n \in \mathbb{N}$. Consequently, (1.26) can be replaced with the following simpler equation:

$$\rho v_{-1} = \max_{x_{-1}, a_{-1}} \left\{ -\omega^* G\left(x_{-1}\right) + \left[(1 - a_{-1}) x_{-1} + \eta\right] \left[v_0 - v_{-1}\right] + a_{-1} x_{-1} \left[v_1 - v_{-1} - \zeta\right] \right\}, \quad (1.36)$$

implying optimal R&D decisions for all followers of the form

$$x_{-1}^* = \max \left\{ G'^{-1} \left(\frac{\max \langle \left[v_0 - v_{-1}\right], \left[v_1 - v_{-1} - \zeta\right] \rangle}{\omega^*} \right), 0 \right\}. \quad (1.37)$$

Let us denote the sequence of value functions under uniform IPR as $\{v_n\}_{n=-1}^{\infty}$. We next establish the existence of a steady-state equilibrium under uniform IPR and characterize some of its most important properties. Establishing the existence of a steady-state equilibrium in this economy is made complicated by the fact that the equilibrium allocation cannot be represented as a solution to a maximization problem. Instead, as emphasized by Definition 1.3, each firm maximizes its value taking the R&D decisions of other firms as given; thus an equilibrium corresponds to a set of R&D decisions that are best responses to themselves and a labor share (wage rate) ω^* that clears the labor market. Nevertheless, there is sufficient structure in the model to guarantee the existence of a steady-state equilibrium and monotonic behavior of values and R&D decisions.

Proposition 1.2 *Consider a uniform IPR policy* $< \eta^{uni}, \zeta^{uni} >$ *and suppose that* $G'^{-1} \left(\left(1 - \lambda^{-1}\right) / (\rho + \eta)\right) > 0$. *Then a steady-state equilibrium* $< \mu^*, \mathbf{v}, a_{-1}^*, \mathbf{x}^*, \omega^*, g^* >$ *exists. Moreover, in any steady-state equilibrium* $\omega^* < 1$. *In addition, if either* $\eta > 0$ *or* $x_{-1}^* > 0$, *then* $g^* > 0$. *For any steady-state R&D decisions* $< a_{-1}^*, \mathbf{x}^* >$, *the steady-state distribution of industries* μ^* *is uniquely determined.*

In addition, we have the following results:

- $v_{-1} \le v_0$ *and* $\{v_n\}_{n=0}^{\infty}$ *forms a bounded and strictly increasing sequence converging to some* $v_\infty \in (0, \infty)$.

- $x_0^* > x_1^*$, $x_0^* \ge x_{-1}^*$, *and* $x_{n+1}^* \le x_n^*$ *for all* $n \in \mathbb{N}$ *with* $x_{n+1}^* < x_n^*$ *if* $x_n^* > 0$. *Moreover, provided that* $G'^{-1} \left(\left(1 - \lambda^{-1}\right) / (\rho + \eta)\right) > 0$ *and* $\zeta > 0$, $x_0^* > x_{-1}^*$.

33

Proof. See the Appendix. ∎

Remark 1.1 The condition that $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$ ensures that there will be positive R&D in equilibrium. If this condition does not hold, then there exists a trivial steady-state equilibrium in which $x_n^* = 0$ for all $n \in \mathbb{Z}_+$, i.e., an equilibrium in which there is no innovation and thus no growth (this follows from the fact that $x_0^* \geq x_n^*$ for all $n \neq 0$, see the Appendix for more details). Moreover, if $\eta > 0$, then this equilibrium would also involve $\mu_0^* = 1$, so that in every industry two firms with equal costs compete a la Bertrand and charge price equal to marginal cost, leading to zero aggregate profits and a labor share of output equal to 1. The assumption that $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$, on the other hand, is sufficient to rule out $\mu_0^* = 1$ and thus $\omega^* = 1$. If, in addition, the steady-state equilibrium involves some probability of catch-up or innovation by the followers, i.e., either $\eta > 0$ or $x_{-1}^* > 0$, then the growth rate is also strictly positive. A sufficient condition to ensure that $x_{-1}^* > 0$ when $\eta = 0$ is that $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\rho - \zeta\right) > 0$.[20]

In addition to the existence of a steady-state equilibrium with positive growth, Proposition 1.2 shows that the sequence of values $\{v_n\}_{n=0}^{\infty}$ is strictly increasing and converges to some v_∞, and more importantly that $\mathbf{x}^* \equiv \{x_n^*\}_{n=1}^{\infty}$ is a decreasing sequence, which implies that technological leaders that are further ahead undertake less R&D. Intuitively, the benefits of further R&D are decreasing in the technology gap, since greater values of the technology gap translate into smaller increases in the equilibrium markup (recall (1.15)). Moreover, the R&D level of neck and-and-neck firms, x_0^*, is greater than both the R&D level of technological leaders that are one step ahead and technological followers that are one step behind (i.e., $x_0^* > x_1^*$ and $x_0^* \geq x_{-1}^*$). This implies that with uniform policy neck-and-neck industries are "most R&D intensive," while industries with the largest technology gaps are "least R&D intensive". This is the basis of the conjecture mentioned in the Introduction that reducing protection given to technologically advanced leaders might be useful for increasing R&D by bringing them into the neck-and-neck state.

---

[20]To see why this condition is sufficient suppose that $\eta = 0$ and also that $x_{-1}^* = 0$. Then (1.36) immediately implies $v_{-1} = 0$ and (1.24) implies $v_1 \geq \left(1 - \lambda^{-1}\right)/\rho$. Moreover, from (1.37) and the fact that $\omega^* \leq 1$, we have $x_{-1}^* \geq G'^{-1}\left(v_1 - v_{-1} - \zeta\right) \geq G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\rho - \zeta\right)$. Therefore, $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\rho - \zeta\right) > 0$ contradicts the hypothesis that $x_{-1}^* = 0$, and implies $x_{-1}^* > 0$. The reason why $\eta > 0$ can, under some circumstances, contribute to positive growth is related to the composition effect discussed above.

1.3.2 State-Dependent IPR Policy

We now extend the results from the previous section to the environment with state-dependent IPR policy, though results on monotonicity of values and R&D efforts no longer hold.[21]

Proposition 1.3 *Consider the state-dependent IPR policy $< \eta, \zeta >$ and suppose that $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/(\rho + \eta)\right.$*

0. Then a steady-state equilibrium $< \mu^*, \mathbf{v}, a^*_{-1}, \mathbf{x}^*, \omega^*, g^* >$ exists. Moreover, in any steady-state equilibrium $\omega^* < 1$. In addition, if either $\eta_1 > 0$ or $x^*_{-1} > 0$, then $g^* > 0$.

Proof. See the Appendix. ∎

Unfortunately, it is not possible to determine the optimal (welfare- or growth-maximizing) state-dependent IPR policy analytically. For this reason, in Section 1.4, we undertake a quantitative investigation of the form and structure of optimal state-dependent IPR policy using plausible parameter values.

1.3.3 Compulsory Versus Bargained License Fees

The analysis so far has characterized the steady-state equilibrium for a given sequence of license fees ζ. Our interpretation in the next section will be that this sequence of license fees is determined by policy—i.e., these fees correspond to compulsory licensing fees for intellectual property that has been patented. We will thus imagine a world in which once a company patents an innovation, the knowledge embedded in this innovation can be used by its competitors as long as they pay a pre-specified licensing fee. The analysis in the next section will show that such licensing fees will increase growth and welfare.

One may also wish to consider an alternative world in which license fees are determined by bilateral bargaining. To characterize the equilibrium in such a world, one must conduct exactly the same analysis as we have done in this section. In other words, one must first characterize the equilibrium for a given sequence of license fees, and then taking the license fees agreed by other firms as given, one ought to consider the bargaining problem between a leader and a follower. As already noted in footnote 11, regardless of whether license fees are set by policy or

---

[21] This is because IPR policies could be very sharply increasing at some technology gap, making a particular state very unattractive for the leader. For example, we could have $\eta_n = 0$ and $\eta_{n+1} \to \infty$, which would imply that $v_{n+1} - v_n$ is negative.

are bargained, no follower will pay a positive price for a license for production (since this will simply lead to zero ex post profits). The only issue then becomes whether the leader and the follower in an industry can agree on a license fee if such fees are not specified by policy.

The answer to this question depends on the exact bargaining protocol between the two firms and there are *two scenarios*, which have quite different implications. The first scenario corresponds to a situation in which the leader and the follower can write a contract specifying the licensing fee *before* the follower undertakes R&D. In this case, the gain to the follower from licensing would be $v_1 - v_0$ (conditional on success in the innovation), while the loss to the leader would be $v_0 - v_{-1}$. As long as the first quantity is greater than the second (which may not always be the case), some level of ζ would be agreed between the two firms. Once this level is determined, then the equilibrium characterization so far applies with this level of license fee. However, this license fee would be *uniform*, not state dependent, since the gain to the follower and the loss to the leader are independent of the technology gap in the industry. Since our analysis in the next section shows significant gains from state-dependent IPR policy (including licensing fees), some type of compulsory licensing fee policy would improve over this bargained solution.

The potential welfare improvements from compulsory licensing policy are greater in the second scenario, where bargaining over the license fee would be undertaken *after* the innovation of the follower. In this scenario, a bargained price may not even emerge. In particular, the follower would have committed itself to an innovation using the patented knowledge of the leader and would have to bargain after its innovation efforts are sunk. In this case, the value of the license to a follower in an industry with an n-step gap is $v_1 - v_{-n}$ instead of $v_1 - v_0$, because if it cannot obtain a license, the follower would still remain n steps behind the leader. Clearly $v_1 - v_{-n} > v_1 - v_0$, reflecting a form of *holdup* of the follower by the leader. This holdup may imply that there would be no licensing fee that the two parties can agree on (for example, because $v_1 - v_{-n}$ will be typically greater than the ex post loss of the leader, $v_0 - v_{-1}$). This second scenario, which appears to us more likely than the scenario with full contracting on the licensing fee before R&D, significantly reduces the role for bilateral licensing arrangements and implies that the type of compulsory licensing policies considered in the next section may create substantial welfare and growth benefits.

1.4 Optimal IPR Policy: A Quantitative Investigation

In this section, we investigate the implications of various different types of IPR policies on R&D, growth and welfare using numerical computations of the steady-state equilibrium. Our purpose is not to provide a detailed calibration of the model economy but to highlight the broad quantitative characteristics of the model and its implications for optimal IPR policy under plausible parameter values. We focus on welfare-maximizing policy (growth-maximizing policies are discussed below). We will see that the structure of optimal IPR policy and the innovation gains from such policy are relatively invariant to the range of parameter values we consider.

1.4.1 Welfare

Our focus so far has been on steady-state equilibria (mainly because of the very challenging nature of transitional dynamics in this class of models). In our quantitative analysis, we continue to focus on steady states and thus look at steady-state welfare. In a steady-state equilibrium, welfare at time $t = 0$ can be written as

$$
\begin{aligned}
\text{Welfare}\,(0) \;\; &= \;\; \int_0^\infty e^{-\rho t} \ln\left(Y\,(0)\, e^{g^* t}\right) dt \\
&= \;\; \frac{\ln Y\,(0)}{\rho} + \frac{g^*}{\rho^2},
\end{aligned}
\tag{1.38}
$$

where the first-line uses the facts that all output is consumed, utility is logarithmic (recall (1.1)), output and consumption at date $t = 0$ are given by $Y\,(0)$, and in the steady-state equilibrium output grows at the rate g^*. The second line simply evaluates the integral. Next, note that

$$
\begin{aligned}
\ln Y\,(t) \;\; &= \;\; \int_0^1 \ln y\,(j, t)\, dj \\
&= \;\; \int_0^1 \ln\left(\frac{q_{-\imath}\,(j, t)\, Y\,(t)}{w\,(t)}\right) dj \\
&= \;\; \int_0^1 \ln q_{-\imath}\,(j, t)\, dj - \ln \omega\,(t) \\
&= \;\; \ln Q\,(t) - \ln \lambda \left(\sum_{n=0}^\infty n\mu_n\,(t)\right) - \ln \omega\,(t),
\end{aligned}
\tag{1.39}
$$

37

where the first line simply uses the definition in (3.3.2), the second line substitutes for $y(j,t)$ from (1.8), the third line uses the definition of the labor share $\omega(t)$, and the final line uses the definition of $Q(t)$ from (1.20) together with the fact that in the steady state $q_i(j,t) = \lambda^n q_{-i}(j,t)$ in a fraction $\mu_n(t)$ of industries. The expression in (1.39) implies that output simply depends on the quality index, $Q(t)$, the distribution of technology gaps, $\mu(t)$ (because this determines markups), and also on the labor share, $\omega(t)$. In steady-state equilibrium, the distribution of technology gaps and labor share are constant, while output and the quality index grow at the steady-state rate g^*. Therefore, for steady-state comparisons of welfare across economies with different policies, it is sufficient to compare two economies with the same level of $Q(0)$, but with different policies. We can then evaluate steady-state welfare with the distribution of industries given by their steady-state values in the two economies, and output and the quality index growing at the corresponding steady-state growth rates. Expression (1.39) also makes it clear that only the aggregate quality index $Q(0)$ needs to be taken to be the same in the different economies. Given $Q(0)$, the dispersion of industries in terms of the quality levels has no effect on output or welfare (though, clearly, the distribution of industries in terms of technology gaps between leaders and followers, μ, influences the level of markups and output, and thus welfare).

However, note one difficulty with welfare comparisons highlighted by equations (1.38) and (1.39); proportional changes in steady-state welfare due to policy changes will depend on the initial level of $Q(0)$, which is an arbitrary number. Therefore, proportional changes in welfare are not informative (though none of this affects ordinal rankings, thus welfare-maximizing policy is well defined and independent of the level of $Q(0)$). These two expressions also make it clear that changes in steady-state welfare will be the sum of two components: the first is the *growth effect*, given by g^*/ρ^2, whereas the second is due to changes in $\ln\lambda\left(\sum_{n=0}^{\infty} n\mu_n\right)/\rho - \ln\omega(0)$. Since changes in the labor share $\omega(0)$ are largely driven by the distribution of industries, we refer to this as the *distribution effect*. Policies will typically affect both of these quantities. In what follows, we give the welfare rankings of different policies and then report the relative magnitudes of the growth and the distribution effects. This will show that the growth effects will be one or two orders of magnitude greater than the distribution effects and dominate welfare comparisons. So if the reader wishes, he or she may think of the magnitudes of the changes in welfare as given by the proportional changes in growth rates.

1.4.2 Calibration

For our calibration exercise, we take the annual discount rate as 5%, i.e., $\rho_{year} = 0.05$. In all our computations, we work with the monthly equivalent of this discount rate in order to increase precision, but throughout the tables, we convert all numbers to their annual counterparts to facilitate interpretation.

The theoretical analysis considered a general production function for R&D given by (1.9). The empirical literature typically assumes a Cobb-Douglas production function. For example, Kortum (1993) considers a function of the form

$$\text{Innovation}\,(t) = B_0 \exp\left(\kappa t\right) (\text{R\&D inputs})^\gamma, \qquad (1.40)$$

where B_0 is a constant and $\exp\left(\kappa t\right)$ is a trend term, which may depend on general technological trends, a drift in technological opportunities, or changes in general equilibrium prices (such as wages of researchers etc.). The advantage of this form is not only its simplicity, but also the fact that most empirical work estimates a single elasticity for the response of innovation rates to R&D inputs. Consequently, they essentially only give information about the parameter γ in terms of equation (1.40). A low value of γ implies that the R&D production function is more concave. For example, Kortum (1993) reports that estimates of γ vary between 0.1 and 0.6 (see also Pakes and Griliches, 1980, or Hall, Hausman and Griliches, 1988). For these reasons, throughout, we adopt a R&D production function similar to (1.40):

$$x = Bh^\gamma \qquad (1.41)$$

where $B, \gamma > 0$. In terms of our previous notation, equation (1.41) implies that $G\left(x\right) = [x/B]^{\frac{1}{\gamma}} w$, where w is the wage rate in the economy (thus in terms of the above function, it is captured by the $\exp\left(\kappa t\right)$ term).[22] Equation (1.41) does not satisfy the boundary conditions we imposed so far and can be easily modified to do so without affecting any of the results, since

---

[22] More specifically, (1.41) can be alternatively written as

$$\text{Innovation}\,(t) = Bw\,(t)^{-\gamma} (\text{R\&D expenditure})^\gamma,$$

thus would be equivalent to (1.40) as long as the growth of $w\,(t)$ can be approximated by constant rate.

in all numerical exercises only a finite number of states are reached.[23] Following the estimates reported in Kortum (1993), we start with a benchmark value of $\gamma = 0.35$, and then report sensitivity checks for $\gamma = 0.1$ and $\gamma = 0.6$. The other parameter in (1.41), B, is chosen so as to ensure an annual growth rate of approximately 1.9%, i.e., $g^* \simeq 0.019$, in the benchmark economy which features indefinitely-enforced patents and no licensing. This growth rate together with $\rho_{year} = 0.05$ also pins down the annual interest rate as $r_{year} = 0.069$ from equation (1.2).

We choose the value of λ using a reasoning similar to Stokey (1995). Equation (1.35) implies that if the expected duration of time between any two consecutive innovations is about 3 years in an industry, then a growth rate of about 1.9% would require $\lambda = 1.05$.[24] This value is also consistent with the empirical findings of Bloom, Schankerman and Van Reenen (2005).[25] We take $\lambda = 1.05$ as the benchmark value, and check the robustness of the results to $\lambda = 1.01$ and $\lambda = 1.2$ (expected duration of 1 year and 12 years, respectively). Finally, without loss of generality, we normalize labor supply to 1. This completes the determination of all the parameters in the model except the IPR policy.

As noted above, we begin with the full patent protection regime without licensing, i.e., $\langle \boldsymbol{\eta} = \{0, 0, ...\}, \boldsymbol{\zeta} = \{\infty, \infty, ...\}\rangle$.[26] We then compare this to an economy with full patent protection and licensing, i.e., $\langle \boldsymbol{\eta} = \{0, 0, ...\}, \boldsymbol{\zeta} = \{\bar{\zeta}, \bar{\zeta}, ...\}\rangle$, where $\bar{\zeta} = v_1 - v_0$.[27] We move to a

---

[23] For example, we could add a small linear term to the production function for R&D, (1.41), and also make it flat after some level $\bar{h}$. For example, the following generalization of (1.41),

$$x = \min\left\{Bh^\gamma + \varepsilon h; B\bar{h}^\gamma + \varepsilon\bar{h}\right\}$$

for ε small and $\bar{h}$ large, makes no difference to our simulation results.

[24] In particular, in our benchmark parameterization with full protection without licensing, 24% of industries are in the neck-and-neck state. This implies that improvements in the technological capability of the economy is driven by the R&D efforts of the leaders in 76% of the industries and the R&D efforts of both the leaders and the followers in 24% of the industries. Therefore, the growth equation, (1.35), implies that $g \simeq \ln\lambda \times 1.24 \times x$, where x denotes the average frequency of innovation in a given industry. A major innovation on average every three years implies a value of $\lambda \simeq 1.05$.

[25] The production function for the intermediate good, (1.5), can be written as $\log(y(j,t)) = n(j,t)\log(\lambda) + \log(l(j,t))$, where $n(j,t)$ is the number of innovations to date in sector j and represents the "knowledge stock" of this industry. Bloom, Schankerman and Van Reenen (2005) proxy the knowledge stock in an industry by the stock of R&D in that industry and estimate the elasticity of sales with respect to the stock of R&D to be approximately 0.06. In terms of the exercise here, this implies that $\log(\lambda) = 0.06$, or that $\lambda \approx 1.06$.

[26] Here $\zeta = \infty$ stands for ζ sufficiently large so that there is no licensing. It does not need to be literally equal to infinity and in fact, in the theoretical analysis we presumed that it is equal to some finite number.

[27] For the interpretation of full patent protection as $\bar{\zeta} = v_1 - v_0$, recall the discussion in footnote 15. Note also that at a license fee of $\bar{\zeta}$, followers are indifferent between $a = 0$ and $a = 1$, and in computing the equilibrium in this case we always suppose that they choose $a = 1$. Thus alternatively one might wish to think that $\bar{\zeta} = v_1 - v_0 - \varepsilon$ for $\varepsilon \downarrow 0$.

comparison of the optimal (welfare-maximizing) uniform IPR policy η^{uni}, ζ^{uni} to the optimal state-dependent IPR policy. Since it is computationally impossible to calculate the optimal value of each η_n and ζ_n, we limit our investigation to a particular form of state-dependent IPR policy, whereby the same η and ζ applies to all industries that have a technology gap of $n = 5$ or more. In other words, the IPR policy can be represented as:

IPR policy $\rightarrow$	none	(η_1, ζ_1)	(η_2, ζ_2)	(η_3, ζ_3)	(η_4, ζ_4)	(η_5, ζ_5)							
Technology gap: $n \rightarrow$	0	1	2	3	4	5	6	7	8	9	10	11 . . ∞

We checked and verified that allowing for further flexibility (e.g., allowing η_5 and η_6 or ζ_5 and ζ_6 to differ) has little effect on our results.

The numerical methodology we pursue relies on *uniformization* and value function iteration. The details of the uniformization technique are described in the proof of Lemma 1.1 in the Appendix. On value function iteration, see Judd (1999). In particular, we first take the IPR policies η and ζ as given and make an initial guess for the equilibrium labor share ω^*. Then for a given ω^*, we generate a sequence of values $\{v_n\}_{n=-\infty}^{\infty}$, and we derive the optimal R&D policies, $\{x_n^*\}_{n=-\infty}^{\infty}$, $\{a_n^*\}_{n=-\infty}^{-1}$ and the steady-state distribution of industries, $\{\mu_n^*\}_{n=0}^{\infty}$. After convergence, we compute the growth rate g^* and welfare, and then check for market clearing in the labor market from equation (1.18). Depending on whether there is excess demand or supply of labor, ω^* is varied and the whole process is repeated until the entire steady-state equilibrium for a given IPR policy is computed. The process is then repeated for different IPR policies.

In the state-dependent IPR case, the optimal (welfare-maximizing) IPR policy sequences, η and ζ, are computed one element at a time, until we find the welfare-maximizing value for that component, for example, η_1. We then move the next component, for example, η_2. Once the welfare-maximizing value of η_2 is determined, we go back to optimize over η_1 again, and this procedure is repeated recursively until convergence.

1.4.3 Full IPR Protection Without Licensing

We start with the benchmark with full protection and no licensing, which is the case that the existing literature has considered so far (e.g., Aghion, Harris, Howitt and Vickers, 2001). In

terms of our model, this corresponds to $\eta_n = 0$ for all n and $\zeta_n = \infty$ for all n. Equation (1.27) implies that $a^*_{-n} = 0$ for all n. We choose the parameter B in terms of (1.41), so that the benchmark economy has an annual growth rate of 1.86%.

The value function for this benchmark case is shown in Figure 1 (with the solid line). The value function has decreasing differences for $n \geq 0$, which is consistent with the results in Proposition 1.2, and features a constant level for all followers (since there is no state dependence in the IPR policy). Figure 2 shows the level of R&D efforts for leaders and followers in this benchmark (with the solid lines). Again consistent with Proposition 1.2, this figure also shows that the R&D level of a leader declines as the technology gap increases and that the highest level of R&D is for firms that are neck-and-neck (i.e., at the technology gap of $n = 0$). Since there is no state-dependent IPR policy, all followers undertake the same level of R&D effort, which is also shown in the figure.

Figure 3 shows the distribution of industries according to technology gaps (again the solid line refers to the benchmark case). The mode of the distribution is at the technology gap of $n = 1$, but there is also a concentration of industries at technology gap $n = 0$, because $a^*_{-n} = 0$ implies that innovations by the followers take them to the "neck-and-neck" state.

The first column of Table 1 also reports the results for this benchmark simulation. As noted above, B is chosen such that the annual growth rate is equal to 0.0186, which is recorded at the bottom of Table 1 together with the initial consumption and welfare levels according to (3.3.2) and (1.38) respectively. The table also shows the R&D levels x^*_0 and x^*_{-1} (0.35 versus 0.22), the frequencies of industries with technology gaps of 0, 1 and 2. The steady-state value of ω is 0.95. Since labor is the only factor of production in the economy, ω^* should not be thought of as the labor share in GDP. Instead, $1 - \omega^*$ measures the share of pure monopoly profits in value added. In the benchmark parameterization, this corresponds to 5% of GDP, which is reasonable.[28] Finally, the table also shows that in this benchmark parameterization 3.2% of the workforce is working as researchers, which is also consistent with US data.[29] These results are encouraging for our simple calibration exercise, since with very few parameter choices, the

---

[28] Bureau of Economic Analysis (2004) reports that the ratio of before-tax profits to GDP in the US economy in 2001 was 7% and the after-tax ratio was 5%.

[29] According to National Science Foundation (2006), the ratio of scientists and engineers in the US workforce in 2001 is about 4%.

model generates reasonable numbers, especially for the share of the workforce allocated to research.[30]

1.4.4 Full IPR Protection With Licensing

We next turn to full IPR protection with licensing. As specified above, we think of full IPR protection with licensing as corresponding to $\eta_n = 0$ for all n (so that patents never expire) and $\zeta_n = \bar{\zeta} \equiv v_1 - v_0$ for all n (so that the license fee for making use of a leading-edge technology is equal to the net present discounted value gap between being a one step ahead leader and a neck-and-neck firm). Figures 1-3 show the corresponding value functions, R&D effort levels and distribution of industries for this case (with the dashed lines). Since there is no state-dependent policy, the general pattern is similar to that in the economy without licensing. There is no longer a spike in R&D effort at $n = 0$, however, since now firms always prefer to pay the license fee and jump ahead of the leading-edge technology. This makes the neck-and-neck state no longer special (in fact, as column 2 of Table 1 shows, in equilibrium there will be no industries in the neck-and-neck state). More importantly, the level of R&D by followers is considerably higher than in the benchmark case. In particular, x^*_{-1} is now 0.25 rather than 0.22. The resulting growth rate is 2.58% instead of 1.86%. Correspondingly, welfare increases by 5.76 points because of the growth effect (that is, g^*/ρ^2 increases by 5.76) and declines by 0.03 points because of the distribution effect (in particular, because of the change in the composition of markups). This case therefore illustrates the general pattern mentioned above, whereby the growth effect is one or two orders of magnitudes greater than the distribution effect and dominates the welfare implications of alternative policies.

It is important to note that, in this case, the boost to growth and welfare comes not from increased R&D effort, but from the fact that the R&D of the followers now also advances the technological frontier of the economy owing to licensing (recall equation (1.35)). In fact, column 2 of Table 1 shows that this considerably higher growth rate is achieved with a *lower* fraction of the workforce, only 2.6%, working in the research sector.

---

[30]Most endogenous growth models imply that a significantly greater fraction of the labor force should be employed in the research sector and one needs to introduce various additional factors to reduce the profitability of research or to make entry into research more difficult. In the current model, the step-by-step nature of innovation and competition plays this role and generates a plausible allocation of workers between research and production.

The contribution of licensing to growth and welfare, which is robust across different parameterizations of the model, is the first important implication of our analysis. Relative to existing models of step-by-step innovation, such as Aghion, Harris, Howitt and Vickers (2001), which do not allow for the possibility of licensing, here the R&D effort by followers can directly contribute to economic growth and this increases the equilibrium growth rate of the economy.

1.4.5 Optimal Uniform IPR Protection

We next turn to optimal IPR policy with licensing. That is, we impose that $\eta_n = \eta$ and $\zeta_n = \zeta$ for all n, and look for values of η and ζ that maximize the welfare in the economy. Column 3 of Table 1 shows that the welfare-maximizing values of η and ζ are both equal to 0 in the benchmark parameterization. This corresponds to zero license fees and indefinite duration of patents, so that followers can never copy the leading-edge technology without R&D, but they can always advance one step ahead of the leader when they are successful in their R&D efforts (without paying any license fees).

The resulting value function, R&D effort levels and industry distributions according to technology gaps are shown in Figures 4-6 (with the solid lines). The figures and column 3 of Table 1 show that the welfare-maximizing IPR policy discourages leaders (this can be seen from the fact that $v_1 - v_0$ declines significantly), but encourages R&D effort by the followers, since when successful they do not have to pay the license fee. The optimal uniform IPR increases the growth rate by only a small amount, however. While the growth rate of the economy with full IPR protection with licensing was 2.58%, it is now 2.63%. This increase in the growth rate also raises steady-state welfare. In particular, the growth effect increases welfare by 4 points, while in this case there is also a slight improvement in welfare because of the change in the distribution of markups (though this is again small, equivalent to 0.1 points, that is, 1/40th as important as the growth effect). Finally, optimal uniform IPR protection also lead to a modest rise in the share of the labor force working in research (from 2.6% to 2.7%).

1.4.6 Optimal State-Dependent IPR Without Licensing

We next turn to our second major question; whether state-dependent IPR makes a significant difference relative to the uniform IPR. To highlight the roles played by different components of

IPR policy, we first investigate the nature of welfare-maximizing space-dependent IPR policy without licensing (so that the comparison is to the benchmark case in column 1). In particular, we set $\zeta_n = \infty$ for all n and look for the combination of $\{\eta_1, ..., \eta_5\}$ that maximizes the welfare. The results are shown in column 4 of Table 1.

Two features are worth noting. First, the growth rate increases noticeably relative to column 1; it is now 2.04% instead of 1.86%. Nevertheless, this increase is models at relative to the benefits of licensing. The increase in steady-state welfare is also correspondingly smaller. Therefore, state-dependent IPR policy with no licensing is not a substitute for licensing.

Second, we see an interesting pattern (which is in fact quite general in all of our quantitative investigations). The optimal state-dependent policy $\{\eta_1, ..., \eta_5\}$ provides *greater* protection to technological leaders that are further ahead. In particular, we find that the optimal policy involves $\eta_1 = 0.71$, $\eta_2 = 0.08$, and $\eta_3 = \eta_4 = \eta_5 = 0$. This corresponds to very little patent protection for firms that are one step ahead of the followers. In particular, since $\eta_1 = 0.71$ and $x^*_{-1} = 0.12$, in this equilibrium firms that are one step behind followers are more than *six times* as likely to catch up with the technological leader because of the expiration of the patent of the leader as they are likely to catch up because of their own successful R&D. Then, there is a steep increase in the protection provided to technological leaders that are two steps ahead, and η_2 is 1/12th of η_1. Perhaps even more remarkably, after a technology gap of three or more steps, optimal IPR involves full protection, and patents never expire.

This pattern of greater protection for technological leaders that are further ahead may go against a naïve intuition that state-dependent IPR policy should try to boost the growth rate of the economy by bringing more industries with large technology gaps (where leaders engage in little R&D) into neck-and-neck competition. This composition effect is present, but dominated by another, more powerful force, the *trickle-down* effect. The intuition for the trickle-down effect is as follows: by providing secure patent protection to firms that are three or more steps ahead of their rivals, optimal state-dependent IPR increases the R&D effort of leaders that are one and two steps ahead as well. This is because technological leaders that are only one or two steps ahead now face greater returns to R&D, which will not only increase their profits but also the security of their intellectual property. Mechanically, high levels of η_1 and η_2 reduce v_1 and v_2, while high IPR protection for more advanced firms increases v_n for $n \geq 3$, and this increases

45

the R&D incentives of leaders at $n = 1$ or at $n = 2$. This pattern of increased R&D investments under state-dependent IPR contrasts with uniform IPR, which always reduces R&D by all firms. The possibility that imperfect state-dependent IPR protection can *increase* (rather than reduce) R&D incentives is a novel feature of our approach and will be illustrated further in the next subsection.

1.4.7 Optimal State-Dependent IPR With Licensing

Finally, we turn to the most general policy regime, which allows both state-dependent patent protection and licensing. In particular, we now choose combinations of $\{\eta_1, ..., \eta_5\}$ and $\{\zeta_1, ..., \zeta_5\}$ to maximize steady-state welfare. The results of this exercise are shown in column 5 of Table 1. The most natural comparison in this case is to the optimal uniform IPR policy with licensing in column 3, where uniform IPR policies η and ζ were chosen to maximize welfare. The value functions, R&D efforts and the industry distribution over different levels of technology gaps in this economy are shown in Figures 4-6 (with the dashed lines).

We see in column 5 that welfare-maximizing IPR policy involves $\eta_n = 0$ for all n, so that with compulsory licensing, optimal IPR involves infinite duration of patents (though this is not always the case, see Table 2). Nevertheless, IPR protection for technological leaders is not full. In particular, the welfare-maximizing policy involves $\zeta_1 = 0$, which implies that followers can build on the leading-edge technology that is one step ahead of their own knowledge without paying any license fees. From there on, ζ increases to $\zeta_2 = 0.98$, then to $\zeta_3 = 1.93$, and to $\zeta_4 = 1.97$. After five steps, the welfare-maximizing policy is equivalent to full patent protection, that is, $\zeta_5 = 1.98$ (note that $v_1 - v_0 = 1.98$). The resulting growth rate of the economy is 2.96%, which is significantly higher than the growth rate under uniform IPR policy, 2.63% in column 3. Steady-state welfare also increases by a corresponding amount relative to the case with optimal uniform IPR case. In particular, the growth effect on welfare is an increase of 1.32 points, while the distribution effect involves a slight deterioration in welfare, equivalent to 0.015 points. Overall, this benchmark case shows that state-dependent policies can increase growth and welfare significantly.

State-dependent policies again achieve this superior growth performance by exploiting the trickle-down effect, which we already saw in the case without licensing. In particular, ζ_n is

an increasing sequence, so that technological leaders that are further ahead receive greater protection. As in the previous subsection, this pattern of IPR is used as a way of boosting the R&D effort of technological leaders that are one or two steps ahead of their rivals (see Figure 5). Since these leaders receive little protection and understand that they can increase both their profits and their IPR protection by undertaking further innovations, they have relatively strong innovation incentives and undertake high levels of R&D. Figure 5 makes it clear that state-dependent relaxation of IPR in this case increases total R&D in the economy relative to full protection. The dashed line in Figure 5 is almost everywhere above the solid line. Alternatively, Table 1 shows that the fraction of the labor force working in R&D increases to 3.9% from 2.6% under full IPR protection with licensing. This positive effect of relaxation of IPR on R&D incentives is a novel implication of our model, and is due to the trickle-down effect.

It is also worth noting that, under state-dependent IPR policy with licensing, the growth rate of the economy receives a further boost from the R&D effort of the followers, since, thanks to licensing, followers' R&D directly contributes to the advancing the technological frontier of the economy. Figure 5 shows that followers that are one step behind the frontier also have a higher R&D effort than even in the case with welfare-maximizing uniform IPR (which involved $\zeta_n = 0$ for all n). The reason for this pattern of R&D efforts is again the trickle-down effect, which increases the value of being a technological leader and thus the incentive of followers to undertake R&D. In contrast, the R&D level of followers that are more than one step behind is lower than in the economy with uniform IPR (though as the comparison of the fraction of the labor force working in research to other columns demonstrates, this is dominated by the increase in the R&D of the technological leaders and of followers that are one step behind).

Overall, the results show that state-dependent IPR policies can increase growth and steady-state welfare substantially, and that this is because of the trickle-down effect. The trickle-down effect is powerful, not only when we consider an economy without licensing, but also in the presence of licensing.

1.4.8 Robustness

Tables 2-5 show the robustness of the patterns documented in Figures 1-6 and in Table 1. In particular, each of these tables changes one of the two parameters λ and γ (increasing or

47

reducing λ to 1.2 or 1.01, and increasing or reducing γ to 0.6 or 0.1) and shows the results corresponding to each one of the five different policy regimes and discussed so far. In each case, we also change the parameter B in equation (1.41) to ensure the growth rate of the benchmark economy with full IPR protection and without licensing is the same as in Table 1, that is, $g^* = 1.86\%$.

Notably, the qualitative, and even the quantitative, patterns in Table 1 are relatively robust. In all cases we see a significant increase in the growth rate and welfare when we allow licensing. The smallest increase is seen when $\gamma = 0.6$, presumably because with limited diminishing returns to R&D, incentives were already sufficiently strong without licensing. As a result, in this case, the growth rate increases only from 1.86% to 1.98%. In all other cases, allowing for licensing increases the growth rate to above 2.6%, which is a sizable increase relative to the baseline of 1.86%.

Moreover, in all cases, moving to state-dependent IPR policy increases the growth rate and welfare further, though the extent of the increase varies depending on parameters.

Perhaps, more noteworthy is the fact that in all cases, welfare-maximizing state-dependent IPR is shaped by the trickle-down effect. In all of the various parameterizations we have considered, there is little or no protection provided to technological leaders that are one step ahead, but IPR protection grows as the technology gap increases. This is the typical pattern implied by the trickle-down effect. In addition, in most, but not all, cases optimal IPR policy provides patents of infinite duration and only makes compulsory licensing fees state dependent. Table 2 and 5, which are for ($\lambda = 1.01$, $\gamma = 0.35$) and ($\lambda = 1.05$, $\gamma = 0.6$), provide instances where both the optimal length of patent enforcement and optimal licensing fees are used as part of the welfare-maximizing policy and are both state dependent.

Finally, we have also computed growth-maximizing policies. In all cases, these are very similar to the welfare-maximizing policies, which is not surprising in view of the fact that, as shown above, welfare comparisons are driven by the growth effects.

We therefore conclude that both the substantial benefits of licensing and the benefits of state-dependent policies are robust across different specifications.

1.4.9 Partial Equilibrium Calibration and Further Robustness

The calibration exercises reported in the previous subsections show that for a range of plausible parameters the trickle-down effect is powerful and induces a pattern of welfare-maximizing (and growth-maximizing) policy that provides greater IPR protection to firms that are technologically more advanced relative to their rivals than to those enjoying a more limited technology gap. This is a new and somewhat surprising finding. Despite the robustness exercises, the reader may wonder whether this result holds for a much broader range of parameter values. Given the computationally-intensive nature of the exercises reported so far, it is not possible to compute or report results for the entire range of parameters for λ, γ and ρ.

In this last subsection, we specialize the economy in three ways and report results for the entire set of parameter values. First, we fix ω, so that the general equilibrium feedback on the labor share is removed. Second, we take the function $G(\cdot)$ to be quadratic. Finally, as assumed by a number of papers in this literature (e.g., Aghion, Harris, Howitt and Vickers, 2001, or Aghion, Bloom, Blundell, Griffith and Howitt, 2005), we assume that the maximum technology gap between a leader and a follower is $n = 2$. Under these assumptions, there are only two possible values for η, η_1 and η_2, and two possible values for ζ, ζ_1 and ζ_2. State-dependent IPR policy here simply means $\eta_1 \neq \eta_2$ and/or $\zeta_1 \neq \zeta_2$. The pattern we have seen in the general equilibrium model, where technological leaders that are further ahead receive greater protection, in turn, corresponds to $\eta_1 > \eta_2$ and/or $\zeta_1 < \zeta_2$. Since there are only two parameters, ρ and λ, we can plot the distribution of optimal policies for a large range of values of these two parameters and see the robust patterns in the form of optimal IPR policy.

More specifically, for a policy vector $(\eta_1, \eta_2, \zeta_1, \zeta_2)$, the stationary equilibrium is character-

49

ized by the solution to the following set of recursive equations:

$$\rho v_2 = \max_{x_2 \geq 0} \left\{ \begin{array}{c} \left(1 - \lambda^{-2}\right) + \left[\left(1 - a^*_{-2}\right) x^*_{-2} + \eta_2\right] [v_0 - v_2] \\ + a^*_{-2} x^*_{-2} [v_{-1} - v_2 + \zeta_2] \end{array} \right\},$$

$$\rho v_1 = \max_{x_1 \geq 0} \left\{ \begin{array}{c} \left(1 - \lambda^{-1}\right) - x_1^2/2 + x_1 [v_2 - v_1] \\ + \left[\left(1 - a^*_{-1}\right) x^*_{-1} + \eta_1\right] [v_0 - v_1] + a^*_{-1} x^*_{-1} [v_{-1} - v_1 + \zeta_1] \end{array} \right\},$$

$$\rho v_0 = \max_{x_0 \geq 0} \left\{ -x_0^2/2 + x_0 [v_1 - v_0] + x_0^* [v_{-1} - v_0] \right\},$$

$$\rho v_{-1} = \max_{x_{-1} \geq 0, a_{-1} \in [0,1]} \left\{ \begin{array}{c} -x_{-1}^2/2 + \left[(1 - a_{-1}) x_{-1} + \eta_1\right] [v_0 - v_{-1}] \\ + a_{-1} x_{-1} [v_1 - v_{-1} - \zeta_1] + x_1^* [v_0 - v_{-1}] \end{array} \right\},$$

and

$$\rho v_{-2} = \max_{x_{-2} \geq 0, a_{-2} \in [0,1]} \left\{ \begin{array}{c} -x_{-2}^2/2 + \left[(1 - a_{-2}) x_{-2} + \eta_2\right] [v_0 - v_{-2}] \\ + a_{-2} x_{-2} [v_1 - v_{-2} - \zeta_2] \end{array} \right\}.$$

Given the solution to these equations, we can determine the welfare-maximizing combination of policies as in our previous calibration exercise (using the same notion of steady-state welfare). For expositional convenience, we do this in two steps, depicted in Figures 7 and 8; first for η_1 and η_2 (setting $\zeta_1 = \zeta_2 = \infty$), and then for ζ_1 and ζ_2 (setting $\eta_1 = \eta_2 = 0$).

Figure 7 shows the pattern of welfare-maximizing policy for the range of parameters $\rho \in [0, 0.5]$ and $\lambda \in (1, 10]$. We can see that for all parameters $\eta_1 > \eta_2$. Thus there is always greater protection given to technological leaders that are two steps ahead than those that are only one step ahead. Figure 8 shows the pattern of optimal policies with only licensing fees for the range of parameters $\rho \in [0, 0.5]$ and $\lambda \in (1, 10]$. Once again, there is greater protection for technological leaders that are further ahead. In fact, in this case for all parameter values, the welfare-maximizing policy involves $\zeta_1 = 0$, meaning that there is no protection provided to technological leaders that are one step ahead. In contrast, ζ_2 is always strictly positive. This pattern again induces a greater R&D investment by technological leaders that are one step ahead of their rivals.

Finally, we have also computed welfare-maximizing policies when the entire vector $(\eta_1, \eta_2, \zeta_1, \zeta_2)$ is allowed to vary. In this case, the welfare-maximizing policy again always provides greater protection to firms that are further ahead. In addition, it typically makes greater use of license fees, but for a small range of parameters, both license fees and relaxation of patent protection

are used simultaneously. To save space, we do not show these results, which are more difficult to depict in the figures.

Overall, these results illustrate that the patterns we found for a narrower range of parameters in the general equilibrium model hold more broadly in this partial equilibrium version of the model. In all cases, there is greater protection given to firms that are further ahead of their rivals, and in all cases, the reason for this is the trickle-down effect.

1.5 Conclusions

In this chapter, we developed a general equilibrium framework to investigate the impact of the extent and form of intellectual property rights (IPR) policy on economic growth and welfare. The two major questions we focused on are whether licensing, which allows followers to build on the leading-edge technology in return of a license fee, has a major impact on the equilibrium growth rate and whether the same degree of patent protection should be given to companies that are further ahead of their competitors as those that are technologically close to their rivals.

In our model economy, firms engage in cumulative (step-by-step) innovation. Leaders can innovate in order to widen the technology gap between themselves and the followers, which enables them to charge higher markups. Followers innovate to catch up with or surpass the technological leaders in their industry. Followers can advance in three different ways. First, the patent of the technological leader may expire, allowing the follower in the industry to copy the leading-edge technology. Second, each follower can undertake "catch-up R&D" to improve its own variant of the product to catch up with the leader. Third, each follower can undertake "frontier R&D," building on and improving the leading-edge technology. In this latter case, when successful, a follower may have to pay a license fee to the technological leader.

In the model economy, IPR policy regulates the length of patents and whether licensing is possible and the cost of licensing. We characterized the form of the steady-state equilibrium and proved its existence under general IPR policies. We then used this framework to investigate the form of "optimal" (welfare-maximizing) IPR policy quantitatively.

The major findings of this quantitative exercise are as follows:

1. A move from an IPR policy without licensing to one that allows for licensing has a

significant effect on the equilibrium growth rate and the welfare. For the benchmark parameterization of our model, licensing increases the growth rate from 1.86% to 2.58% per annum, which is a significant effect. There is a corresponding increase in welfare as well. These substantial increases are robust to a large range of variation in the parameters.

2. State-dependent IPR also leads to a significant improvement in the equilibrium growth rate and welfare. In our benchmark parameterization, welfare-maximizing IPR policy increases the growth rate of the economy from 2.58% under the best possible uniform IPR policy to 2.96% under state-dependent IPR policy. Perhaps more interesting than this substantial impact on both growth and welfare is the form of the optimal state-dependent IPR policy. Contrary to a naïve intuition, we find that the welfare-maximizing IPR policy provides greater protection to firms that are further ahead of their rivals than those that are technologically close to their competitors. Underlying this form of the optimal IPR policy is *the trickle-down effect*. The trickle-down effect implies that providing greater protection to sufficiently advanced technological leaders not only increases their R&D efforts but also raises the R&D efforts of all technological leaders that are less advanced than this level. This is because the reward to innovation now includes the greater protection that they will receive once they reach this higher level of technology. Our results suggest that the trickle-down effect is powerful both with and without licensing, and its form and magnitude are relatively insensitive to the exact parameter values used in the quantitative investigation.

The analysis in this chapter suggests that a move to a richer menu of IPR policies, in particular, a move towards optimal state-dependent policies with licensing, may significantly increase innovation, economic growth and welfare. The results also show that the form of optimal IPR policy may depend on the industry structure (and the technology of catch-up within the industry). It should be noted, however, that these conclusions are based on a quantitative evaluation of a rather simple model. Our objective has not been to obtain practical policy prescriptions and the exact effects of different policies implied by our model undoubtedly miss a host of important factors and ignore potential limitations on the form and complexity of IPR policies. Nevertheless, our results demonstrate a range of robust and new effects that

should be part of the calculus of IPR and competition policy.

The next step in this line of research should be to investigate the robustness of these effects in different models of industry dynamics. It would also be useful to study whether the relationship between the form of optimal IPR policy and industry structure suggested by our analysis also applies when variation in industry structure has other sources (for example, differences in the extent of fixed costs causing differential gaps between technological leaders and followers across industries). The most important area for future work is a detailed empirical investigation of the form of optimal IPR policy, using both better estimates of the effects of IPR policy on innovation rates and also structural models that would enable the evaluation of the effects of different policies on equilibrium growth and welfare. We hope that the theoretical framework presented in this chapter will be useful in developing models that can be estimated in future work.

1.6 Appendix: Proofs

Derivation of Equation (1.22)

Fix the equilibrium R&D policies of other firms, $x^*_{-n}(t)$ and $a^*_{-n}(t)$, the equilibrium interest and wage rates, $r^*(t)$ and $w^*(t)$, and equilibrium profits $\{\Pi^*_n(t)\}^{\infty}_{n=1}$. Then the value of the firm that is n steps ahead at time t can be written as:

$$
V_n(t) = \max_{x_n(t)} \{[\Pi^*_n(t) - w^*(t) G(x_n(t))] \Delta t + o(\Delta t) \tag{1.42}
$$

$$
+ \exp(-r^*(t+\Delta t)\Delta t)
\begin{bmatrix}
(x_n(t)\Delta t + o(\Delta t)) V_{n+1}(t+\Delta t) \\[2mm]
+ (\eta_n \Delta t + (1 - a^*_{-n}(t)) x^*_{-n}(t)\Delta t + o(\Delta t)) V_0(t+\Delta t) \\[2mm]
+ (a^*_{-n}(t) x^*_{-n}(t)\Delta t + o(\Delta t)) (V_{-1}(t+\Delta t) + \hat{\zeta}_n) \\[2mm]
+ (1 - x_n(t)\Delta t - \eta_n \Delta t - x^*_{-n}(t)\Delta t - o(\Delta t)) V_n(t+\Delta t)
\end{bmatrix}
\}.
$$

The first part of this expression is the flow profits minus R&D expenditures during a time interval of length Δt. The second part is the continuation value after this interval has elapsed. $V_{n+1}(t)$ and $V_0(t)$ are defined as net present discounted values for a leader that is $n+1$ steps ahead and a firm in an industry that is neck-and-neck (i.e., $n = 0$). The second part of the expression uses the fact that in a short time interval Δt, the probability of innovation by the leader is $x_n(t)\Delta t + o(\Delta t)$, where $o(\Delta t)$ again denotes second-order terms. This explains the first line of the continuation value. For the remainder of the continuation value, note that the probability that the follower will catch up with the leader is $(1 - a^*_{-n}(t)) x^*_{-n}(t)\Delta t + o(\Delta t)$; in particular, if $a^*_{-n}(t) = 1$, this eventually will never happen, since the follower would be undertaking R&D not to catch up but to surpass the leader. This explains the third line, which applies when $a^*_{-n}(t) = 1$. There are two differences between the second and third lines; (i) in the third line, conditional on success by the follower, a leader moves to the position of a follower rather than a neck-and-neck firm (V_{-1} instead of V_0); (ii) it receives the state-dependent patent fee $\hat{\zeta}_n$. Finally, the last line applies when no R&D effort is successful and patents continue to be enforced, so that the technology gap remains at n steps. Now, subtract $V_n(t)$ from both sides, divide everything by Δt, and take the limit as $\Delta t \to 0$ to obtain (1.22). ∎

Proof of Proposition 1.1

Equations (1.19) and (1.21) imply

$$
Y(t) = \frac{w(t)}{\omega(t)} = \frac{Q(t)\lambda^{-\sum^{\infty}_{n=0} n\mu^*_n(t)}}{\omega(t)}.
$$

Since $\omega(t) = \omega^*$ and $\{\mu^*_n\}^{\infty}_{n=0}$ are constant in steady state, $Y(t)$ grows at the same rate as $Q(t)$. Therefore,

$$
g^* = \lim_{\Delta t \to 0} \frac{\ln Q(t+\Delta t) - \ln Q(t)}{\Delta t}.
$$

Now note the following: during an interval of length Δt (i) in the fraction μ^*_n of the industries with technology gap $n \geq 1$ the leaders innovate at a rate $x^*_n \Delta t + o(\Delta t)$; (ii) in the same industries, the followers innovate at the rate $a^*_{-n} x^*_{-n} \Delta t + o(\Delta t)$; (iii) in the fraction μ^*_0 of the industries with technology gap of $n = 0$, both firms innovate, so that the total innovation rate is $2x^*_0 \Delta t + o(\Delta t)$); and (iv) each

54

innovation increase productivity by a factor λ. Combining these observations, we have

$$\ln Q\left(t + \Delta t\right) = \ln Q\left(t\right) + \ln \lambda \left[2\mu_0^* x_0^* \Delta t + \sum_{n=1}^{\infty} \mu_n^* x_n^* \Delta t + o\left(\Delta t\right) + \sum_{n=1}^{\infty} a_{-n}^* x_{-n}^* \Delta t + o\left(\Delta t\right)\right].$$

Subtracting $\ln Q\left(t\right)$, dividing by Δt and taking the limit $\Delta t \to 0$ gives (1.35). ∎

Proof of Proposition 1.2

We prove this proposition in four parts. (1) Existence of a steady-state equilibrium. (2) Properties of the sequence of value functions. (3) Properties of the sequence of R&D decisions. (4) Uniqueness of an invariant distribution given R&D policies.

Part 1: Existence of a Steady-State Equilibrium.

First, note that each x_n belongs to a compact interval $[0, \bar{x}]$, where $\bar{x}$ is the maximal flow rate of innovation defined in (1.11) above. Now fix a labor share $\tilde{\omega} \in [0, 1]$ and a sequence $\langle \tilde{a}_{-1}, \tilde{\mathbf{x}} \rangle$ of (Markovian) steady-state strategies for all other firms in the economy, and consider the dynamic optimization problem of a single firm. Our first result characterizes this problem and shows that given some $\mathbf{z} \equiv \langle \tilde{\omega}, \tilde{a}_{-1}, \tilde{\mathbf{x}} \rangle$, the value function of an individual firm is uniquely determined, while its optimal R&D choices are given by a convex-valued correspondence. In what follows, we denote sets and correspondences by uppercase letters and refer to their elements by lowercase letters, e.g., $a_{-1}\left(\mathbf{z}\right) \in A_{-1}\left[\mathbf{z}\right]$, $x_n\left(\mathbf{z}\right) \in X_n\left[\mathbf{z}\right]$.

Lemma 1.1 *Consider a uniform IPR policy $\left\langle \eta^{uni}, \zeta^{uni} \right\rangle$, and suppose that the labor share and the R&D policies of all other firms are given by $\mathbf{z} = \langle \tilde{\omega}, \tilde{a}_{-1}, \tilde{\mathbf{x}} \rangle$. Then the dynamic optimization problem of an individual firm leads to a unique value function $\mathbf{v}\left[\mathbf{z}\right] : \{-1\} \cup \mathbb{Z}_+ \to \mathbb{R}_+$ and optimal R&D policies $\hat{A}_{-1}\left[\mathbf{z}\right] \subset [0, 1]$ and $\hat{\mathbf{X}}\left[\mathbf{z}\right] : \{-1\} \cup \mathbb{Z}_+ \rightrightarrows [0, \bar{x}]$ are compact and convex-valued for each $\mathbf{z} \in \mathbf{Z}$ and upper hemi-continuous in $\mathbf{z}$ (where $\mathbf{v}\left[\mathbf{z}\right] \equiv \{v_n\left[\mathbf{z}\right]\}_{n=-1}^{\infty}$ and $\hat{\mathbf{X}}\left[\mathbf{z}\right] \equiv \left\{ \hat{X}_n\left[\mathbf{z}\right] \right\}_{n=-1}^{\infty}$).*

Proof. Fix $\mathbf{z} = \left\langle \tilde{\omega}, \{\tilde{x}_n\}_{n=-1}^{\infty}, \{\tilde{a}_n\}_{n=-\infty}^{-1} \right\rangle$, and consider the optimization problem of a representative firm, written recursively as:

$$\rho v_n = \max_{x_n \in [0, \bar{x}]} \left\{ \left(1 - \lambda^{-n}\right) - \tilde{\omega} G\left(x_n\right) + x_n \left[v_{n+1} - v_n\right] \right.$$
$$+ \quad \tilde{x}_{-1}\left(\tilde{a}_{-1}\left[v_{-1} - v_n + \zeta\right] + \left(1 - \tilde{a}_{-1}\right)\left[v_0 - v_n\right]\right) + \eta \left[v_0 - v_n\right] \right\} \text{ for } n \in \mathbb{N}$$

$$\rho v_0 = \max_{x_0 \in [0, \bar{x}]} \left\{ -\tilde{\omega} G\left(x_0\right) + x_0 \left[v_1 - v_0\right] + \tilde{x}_0 \left[v_{-1} - v_0\right] \right\}$$

$$\rho v_{-1} = \max_{x_{-1} \in [0, \bar{x}], \, a_{-1} \in [0, 1]} \left\{ -\tilde{\omega} G\left(x_0\right) + x_{-1}\left(a_{-1}\left[v_1 - v_{-1} - \zeta\right] + \left(1 - a_{-1}\right)\left[v_0 - v_{-1}\right]\right) \right.$$
$$+ \eta \left[v_0 - v_{-1}\right] \right\}.$$

We now transform this dynamic optimization problem into a form that can be represented as a contraction mapping using the method of "uniformization" (see, for example, Ross, 1996, Chapter 5). Let $\tilde{\boldsymbol{\xi}} = \left\langle \{\tilde{x}_n\}_{n=-1}^{\infty}, \{\tilde{a}_n\}_{n=-\infty}^{-1} \right\rangle$ and $p_{n,n'}\left(\boldsymbol{\xi} \mid \tilde{\boldsymbol{\xi}}\right)$ be the probability that the next state will be n' starting with state n when the firm in question chooses policies $\boldsymbol{\xi} \equiv \left\langle \{x_n\}_{n=-1}^{\infty}, \{a_n\}_{n=-\infty}^{-1} \right\rangle$ and the R&D policy of other firms is given by $\tilde{\boldsymbol{\xi}}$. Using the fact that, because of uniform IPR policy, $\langle x_{-n}, a_{-n} \rangle = \langle x_{-1}, a_{-1} \rangle$ for

all $n \in \mathbb{N}$, these transition probabilities can be written as:

$p_{-1,0}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{(1-a_{-1})x_{-1}+\eta}{x_{-1}+\eta}$	$p_{-1,1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{a_{-1}x_{-1}}{x_{-1}+\eta}$	
$p_{0,-1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{\tilde{x}_0}{x_0+\tilde{x}_0}$	$p_{0,1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{x_0}{x_0+\tilde{x}_0}$	
$p_{n,-1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{a_{-1}\tilde{x}_{-1}}{x_n+\tilde{x}_{-1}+\eta}$	$p_{n,0}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{(1-a_{-1})\tilde{x}_{-1}+\eta}{x_n+\tilde{x}_{-1}+\eta}$	$p_{n,n+1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{x_n}{x_n+\tilde{x}_{-1}+\eta}$

Uniformization involves adding fictitious transitions from a state into itself, which do not change the value of the program, but allow us to represent the optimization problem as a contraction. For this purpose, define the transition rates ψ_n as

$$\psi_n\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\begin{cases} x_n+x_{-1}+\eta & \text{for } n\in\{1,2,...\} \\ x_{-1}+\eta & \text{for } n=-1 \\ 2x_n & \text{for } n=0 \end{cases}.$$

These transition rates are finite since $\psi_n\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)\le\psi\equiv2\bar{x}+\eta<\infty$ for all n, where $\bar{x}$ is the maximal flow rate of innovation defined in (1.11) in the text (both $\bar{x}$ and η are finite by assumption).

Now following equation (5.8.3) in Ross (1996), we can use these transition rates and define the new transition probabilities (including the fictitious transitions from a state to itself) as:

$$\tilde{p}_{n,n'}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\begin{cases} \frac{\psi_n\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)}{\psi}p_{n,n'}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right) & \text{if } n\ne n' \\ 1-\frac{\psi_n\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)}{\psi} & \text{if } n=n' \end{cases}.$$

This yields equivalent transition probabilities

$\tilde{p}_{-1,-1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=1-\frac{x_{-1}+\eta}{2\bar{x}+\eta}$	$\tilde{p}_{-1,0}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{(1-a_{-1})x_{-1}+\eta}{2\bar{x}+\eta}$	$\tilde{p}_{-1,1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{a_{-1}x_{-1}}{2\bar{x}+\eta}$	
$\tilde{p}_{0,-1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{\tilde{x}_0}{2\bar{x}+\eta}$	$\tilde{p}_{0,0}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=1-\frac{x_0+\tilde{x}_0}{2\bar{x}+\eta}$	$\tilde{p}_{0,1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{x_0}{2\bar{x}+\eta}$	
$\tilde{p}_{n,0}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{(1-\tilde{a}_{-1})\tilde{x}_{-1}+\eta}{2\bar{x}+\eta}$	$\tilde{p}_{n,n}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=1-\frac{x_n+\tilde{x}_{-1}+\eta}{2\bar{x}+\eta}$	$\tilde{p}_{n,n+1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{x_n}{2\bar{x}+\eta}$	$\tilde{p}_{n,-1}\left(\boldsymbol{\xi}\mid\tilde{\boldsymbol{\xi}}\right)=\frac{\tilde{a}_{-1}\tilde{x}_{-1}}{2\bar{x}+\eta}$

and also defines an effective discount factor β given by

$$\beta\equiv\frac{\psi}{\rho+\psi}=\frac{2\bar{x}+\eta}{\rho+2\bar{x}+\eta}.$$

Also let the per period return function (profit net of R&D expenditures) be

$$\hat{\Pi}_n\left(x_n\right)=\begin{cases} \frac{1-\lambda^{-n}-\tilde{\omega}G(x_n)}{\rho+2\bar{x}+\eta} & \text{if } n\ge1 \\ \frac{-\tilde{\omega}G(x_n)}{\rho+2\bar{x}+\eta} & \text{otherwise} \end{cases}. \tag{1.43}$$

Using these transformations, the dynamic optimization problem can be written as:

$$\begin{aligned} v_n &= \max_{x_n,\,a_n}\left\{\hat{\Pi}_n\left(x_n\right)+\beta\sum_{n'}\tilde{p}_{n,n'}\left(\xi_n\mid\tilde{\boldsymbol{\xi}}\right)\tilde{v}_{n'}\right\}, \quad\text{for all } n\in\mathbb{Z}, \tag{1.44} \\ &\equiv T\tilde{v}_n, \quad\text{for all } n\in\mathbb{Z}. \end{aligned}$$

where $\mathbf{v}\equiv\{v_n\}_{n=-1}^{\infty}$ and the second line defines the operator T, mapping from the space of functions $\mathbf{V}\equiv\{\mathbf{v}:\{-1\}\cup\mathbb{Z}_+\to\mathbb{R}_+\}$ into itself. T is clearly a contraction, thus, for given $\mathbf{z}=\left\langle\tilde{\omega},\tilde{a}_{-1},\{\tilde{x}_n\}_{n=-1}^{\infty}\right\rangle$, possesses a unique fixed point $\mathbf{v}^*\equiv\{v_n^*\}_{n=-1}^{\infty}$ (e.g., Stokey, Lucas and Prescott, 1989).

56

Moreover, $x_n \in [0, \bar{x}]$, $a_{-1} \in [0, 1]$, and v_n for each $n = -1, 0, 1, \dots$ given by the right-hand side of (1.44) is continuous in a_n and x_n (a_n applying only for $n = -1$), so Berge's Maximum Theorem (Aliprantis and Border, 1999, Theorem 16.31, p. 539) implies that the set of maximizers $\left\langle \hat{A}_{-1}, \left\{ \hat{X}_n \right\}_{n=-1}^{\infty} \right\rangle$ exists, is nonempty and compact-valued for each $\mathbf{z}$ and is upper hemi-continuous in $\mathbf{z} = \left\langle \tilde{\omega}, \tilde{a}_{-1}, \{\tilde{x}_n\}_{n=-1}^{\infty} \right\rangle$. Moreover, concavity of v_n in a_n and x_n for each $n = -1, 0, 1, \dots$ implies that $\left\langle \hat{A}_{-1}, \left\{ \hat{X}_n \right\}_{n=-1}^{\infty} \right\rangle$ is also convex-valued for each $\mathbf{z}$, completing the proof. ∎

Now let us start with an arbitrary $\mathbf{z} \equiv \langle \tilde{\omega}, \tilde{a}_{-1}, \tilde{\mathbf{x}} \rangle \in \mathbf{Z} \equiv [0,1]^2 \times [0, \bar{x}]^{\infty}$. From Lemma 1.1, this $\mathbf{z}$ is mapped into optimal R&D decision sets $\hat{A}_{-1}[\mathbf{z}]$ and $\hat{\mathbf{X}}[\mathbf{z}]$, where $\hat{a}_{-1} \in \hat{A}_{-1}[\mathbf{z}]$ and $\hat{x}_n[\mathbf{z}] \in \hat{X}_n[\mathbf{z}]$. From R&D policies $\langle \tilde{a}_{-1}, \tilde{\mathbf{x}} \rangle$, we calculate $\boldsymbol{\mu}[\tilde{a}_{-1}, \tilde{\mathbf{x}}] \equiv \{\mu_n[\tilde{a}_{-1}, \tilde{\mathbf{x}}]\}_{n=0}^{\infty}$ using equations (1.31), (1.32) and (1.33). Then we can rewrite the labor market clearing condition (1.34) as

$$
\begin{aligned}
\omega &= \min\left\{ \sum_{n=0}^{\infty} \mu_n \left[\frac{1}{\lambda^n} + G(\tilde{x}_n)\tilde{\omega} + G(\tilde{x}_{-n}) \right] \tilde{\omega}; 1 \right\}, \\
&\equiv \varphi(\tilde{\omega}, \tilde{a}_{-1}, \tilde{\mathbf{x}})
\end{aligned}
\tag{1.45}
$$

where due to uniform IPR, $\hat{x}_{-n} = \hat{x}_{-1}$ for all $n > 0$. Next, define the mapping (correspondence)

$$
\Phi[\mathbf{z}] \equiv \left(\varphi(\mathbf{z}), \ \hat{A}_{-1}[\mathbf{z}], \ \hat{\mathbf{X}}[\mathbf{z}] \right)
$$

, which maps $\mathbf{Z}$ into itself, that is,

$$
\Phi \colon \mathbf{Z} \rightrightarrows \mathbf{Z}.
\tag{1.46}
$$

That Φ maps $\mathbf{Z}$ into itself follows since $\mathbf{z} \in \mathbf{Z}$ consists of $\tilde{a}_{-1} \in [0,1]$, $\tilde{\mathbf{x}} \in [0, \bar{x}]^{\infty}$ and $\tilde{\omega} \in [0,1]$, and the image of $\mathbf{z}$ under Φ consists of $\hat{a}_{-1} \in [0,1]$ and $\hat{\mathbf{x}} \in [0, \bar{x}]^{\infty}$, and moreover, (1.45) is clearly in $[0,1]$ (since the right-hand side is nonnegative and bounded above by 1). Finally, from Lemma 1.1, $\hat{A}_{-1}[\mathbf{z}]$ and $\hat{X}_n[\mathbf{z}]$ are compact and convex-valued for each $\mathbf{z} \in \mathbf{Z}$, and also upper hemi-continuous in $\mathbf{z}$, and φ is continuous. Using this construction, we can establish the existence of a steady-state equilibrium as follows.

We first show that the mapping $\Phi \colon \mathbf{Z} \rightrightarrows \mathbf{Z}$ constructed in (1.46) has a fixed point, and then establish that when $G'^{-1}\left(\left(1 - \lambda^{-1}\right) / (\rho + \eta) \right) > 0$ this fixed point corresponds to a steady state with $\omega^* < 1$. First, it has already been established that Φ maps $\mathbf{Z}$ into itself. We next show that $\mathbf{Z}$ is compact in the product topology and is a subset of a locally convex Hausdorff space. The first part follows from the fact that $\mathbf{Z}$ can be written as the Cartesian product of compact subsets, $\mathbf{Z} = [0,1] \times [0,1] \times \prod_{n=-1}^{\infty} [0, \bar{x}]$. Then by Tychonoff's Theorem (e.g., Aliprantis and Border, 1999, Theorem 2.57, p. 52; Kelley, 1955, p. 143), $\mathbf{Z}$ is compact in the product topology. Moreover, $\mathbf{Z}$ is clearly nonempty and also convex, since for any $\mathbf{z}, \mathbf{z}' \in \mathbf{Z}$ and $\lambda \in [0,1]$, we have $\lambda \mathbf{z} + (1 - \lambda)\mathbf{z}' \in \mathbf{Z}$. Finally, since $\mathbf{Z}$ is a product of intervals on the real line, it is a subset of a locally convex Hausdorff space (see Aliprantis and Border, 1999, Lemma 5.54, p. 192).

Next, φ is a continuous function from $\mathbf{Z}$ into $[0,1]$ and from Lemma 1.1, $\hat{A}_{-1}(\mathbf{z})$ and $\hat{X}_n(\mathbf{z})$ for $n \in \{-1\} \cup \mathbb{Z}_+$ are upper hemi-continuous in $\mathbf{z}$. Consequently, $\Phi \equiv \left\langle \varphi[\mathbf{z}], \ \hat{A}_{-1}[\mathbf{z}], \ \hat{\mathbf{X}}[\mathbf{z}] \right\rangle$ has closed graph in $\mathbf{z}$ in the product topology. Moreover, each one of $\varphi(\mathbf{z})$, $\hat{A}_{-1}(\mathbf{z})$ and $\hat{X}_n(\mathbf{z})$ for $n = -1, 0, \dots$ is nonempty, compact and convex-valued. Therefore, the image of the mapping Φ is nonempty, compact and convex-valued for each $\mathbf{z} \in \mathbf{Z}$. The Kakutani-Fan-Glicksberg Fixed Point Theorem implies that if the function Φ maps a convex, compact and nonempty subset of a locally convex Hausdorff space into itself and has closed graph and is nonempty, compact and convex-valued $\mathbf{z}$, then it possesses a fixed point $\mathbf{z}^* \in \Phi(\mathbf{z}^*)$ (see Aliprantis and Border, 1999, Theorem 16.50 and Corollary 16.51, p. 549-550). This establishes the existence of a fixed point $\mathbf{z}^*$ of Φ.

To complete the proof, we need to show that the fixed point, $\mathbf{z}^*$, corresponds to a steady state

equilibrium. First, since $\hat{a}_n \left(\omega^*, a^*_{-1}, \{x^*_n\}^\infty_{n=-1} \right) = a^*_{-1}$ and $\hat{x}_n \left(\omega^*, a^*_{-1}, \{x^*_n\}^\infty_{n=-1} \right) = x^*_n$ for $n \in \{-1\} \cup \mathbb{Z}_+$, we have that given a labor share of ω^*, $\left\langle a^*_{-1}, \{x^*_n\}^\infty_{n=-1} \right\rangle$ constitutes an R&D policy vector that is best response to itself, as required by steady-state equilibrium (Definition 1.3). Next, we need to prove that the implied labor share ω^* leads to labor market clearing. This follows from the fact that the fixed point involves $\omega^* < 1$, since in this case (1.45) will have an interior solution, ensuring labor market clearing. Suppose, to obtain a contradiction, that $\omega^* = 1$. Then, as noted in the text, we must have $\mu^*_0 = 1$. From (1.31), (1.32) and (1.33), this implies $x^*_n = 0$ for $n \in \{-1\} \cup \mathbb{Z}_+$. However, Lemma 1.2 implies that this is not possible when $G'^{-1} \left(\left(1 - \lambda^{-1} \right) / (\rho + \eta) \right) > 0$. Consequently, (1.45) cannot be satisfied at $\omega^* = 1$, implying that $\omega^* < 1$. When $\omega^* < 1$, the labor market clearing condition (1.34) is satisfied at ω^* as an equality, so ω^* is an equilibrium given $\{x^*_n\}^\infty_{n=-1}$, and thus $\mathbf{z}^* = \left(\omega^*, a^*_{-1}, \{x^*_n\}^\infty_{n=-1} \right)$ is a steady-state equilibrium as desired.

Finally, if $\eta > 0$, then (1.33) implies that $\mu^*_0 > 0$. Since $x^*_0 > 0$ from Lemma 1.2, equation (1.35) implies $g^* > 0$. Alternatively, if $x^*_{-1} > 0$, then $g^* > 0$ follows from (1.35). This completes the proof of the existence of a steady-state equilibrium with positive growth.

Part 2: Properties of the Sequence of Value Functions.

Let $\left\langle a_{-1}, \{x_n\}^\infty_{n=-1} \right\rangle$ be the R&D decisions of the firm and $\{v_n\}^\infty_{n=-1}$ be the sequence of values, taking the decisions of other firms and the industry distributions, $\{x^*_n\}^\infty_{n=-1}$, $\{\mu^*_n\}^\infty_{n=-1}$, ω^* and g, as given. By choosing $x_n = 0$ for all $n \geq -1$, the firm guarantees $v_n \geq 0$ for all $n \geq -1$. Moreover, since flow profit satisfy $\pi_n \leq 1$ for all $n \geq -1$, $v_n \leq 1/\rho$ for all $n \geq -1$, establishing that $\{v_n\}^\infty_{n=-1}$ is a bounded sequence, with $v_n \in [0, 1/\rho]$ for all $n \geq -1$.

Proof of $v_1 > v_0$: Suppose, first, $v_1 \leq v_0$, then (1.30) implies $x^*_0 = 0$, and by the symmetry of the problem in equilibrium (1.25) implies $v_0 = v_1 = 0$. As a result, from (1.29) we obtain $x^*_{-1} = 0$. Equation (1.24) implies that when $x^*_{-1} = 0$, $v_1 \geq \left(1 - \lambda^{-1} \right) / (\rho + \eta) > 0$, yielding a contradiction and proving that $v_1 > v_0$. $\square$

Proof of $v_{-1} \leq v_0$: Suppose, to obtain a contradiction, that $v_{-1} > v_0$.

If $v_1 - \zeta \leq v_0$, (1.29) yields $x^*_{-1} = 0$. This implies $v_{-1} = \eta v_0 / (\rho + \eta)$, which contradicts $v_{-1} > v_0$ since $\eta / (\rho + \eta) < 1$. Thus we must have $v_1 - \zeta > v_0$, which implies that $a^*_{-1} = 1$. Imposing $a^*_{-1} = 1$, the value function of a neck-and-neck firm can be written as:

$$
\begin{aligned}
\rho v_0 &= \max_{x_0} \left\{ -\omega^* G(x_0) + x_0 [v_1 - v_0] + x^*_0 [v_{-1} - v_0] \right\}, & (1.47) \\
&\geq \max_{x_0} \left\{ -\omega^* G(x_0) + x_0 [v_1 - v_0] \right\}, \\
&\geq \max_{x_0} \left\{ -\omega^* G(x_0) + x_0 [v_1 - v_{-1} - \zeta] \right\}, \\
&\geq -\omega^* G \left(x^*_{-1} \right) + x^*_{-1} [v_1 - v_{-1} - \zeta], \\
&\geq -\omega^* G \left(x^*_{-1} \right) + x^*_{-1} [v_1 - v_{-1} - \zeta] + \eta [v_0 - v_{-1}], \\
&= \rho v_{-1},
\end{aligned}
$$

which contradicts the hypothesis that $v_{-1} > v_0$ and establishes the claim. $\square$

Proof of $v_n < v_{n+1}$: Suppose, to obtain a contradiction, that $v_n \geq v_{n+1}$. Now (1.28) implies $x^*_n = 0$, and (1.24) becomes

$$
\rho v_n = \left(1 - \lambda^{-n} \right) + x^*_{-1} \left[a^*_{-1} (v_{-1} + \zeta) + \left(1 - a^*_{-1} \right) v_0 - v_n \right] + \eta [v_0 - v_n] \tag{1.48}
$$

Also from (1.24), the value for state $n + 1$ satisfies

$$
\rho v_{n+1} \geq \left(1 - \lambda^{-n-1} \right) + x^*_{-1} \left[a^*_{-1} (v_{-1} + \zeta) + \left(1 - a^*_{-1} \right) v_0 - v_{n+1} \right] + \eta [v_0 - v_{n+1}]. \tag{1.49}
$$

58

Combining the two previous expressions, we obtain

$$\left(1 - \lambda^{-n}\right) + x^*_{-1}\left[a^*_{-1}\left(v_{-1} + \zeta\right) + \left(1 - a^*_{-1}\right)v_0 - v_n\right] + \eta\left[v_0 - v_n\right]$$
$$\geq \quad 1 - \lambda^{-n-1} + x^*_{-1}\left[a^*_{-1}\left(v_{-1} + \zeta\right) + \left(1 - a^*_{-1}\right)v_0 - v_{n+1}\right] + \eta\left[v_0 - v_{n+1}\right].$$

Since $\lambda^{-n-1} < \lambda^{-n}$, this implies $v_n < v_{n+1}$, contradicting the hypothesis that $v_n \geq v_{n+1}$, and establishing the desired result, $v_n < v_{n+1}$. Consequently, $\{v_n\}_{n=-1}^{\infty}$ is nondecreasing and $\{v_n\}_{n=0}^{\infty}$ is (strictly) increasing. Since a nondecreasing sequence in a compact set must converge, $\{v_n\}_{n=-1}^{\infty}$ converges to its limit point, v_∞, which must be strictly positive, since $\{v_n\}_{n=0}^{\infty}$ is strictly increasing and has a nonnegative initial value. $\square$

The above results combined complete the proof that values form an increasing sequence. ∎

Part 3: Properties of the Sequence of R&D Decisions.

Proof of $x^*_{n+1} < x^*_n$: From equation (1.28),

$$\delta_{n+1} \equiv v_{n+1} - v_n < v_n - v_{n-1} \equiv \delta_n \tag{1.50}$$

would be sufficient to establish that $x^*_{n+1} < x^*_n$ whenever $x^*_n > 0$. We next show that this is the case.

Let us write:

$$\bar{\rho}v_n = \max_{x_n}\left\{\left(1 - \lambda^{-n}\right) - \omega^* G\left(x_n\right) + x^*_n\left[v_{n+1} - v_n\right] + x^*_{-1}\left[a^*_{-1}\left(v_{-1} + \zeta\right) + \left(1 - a^*_{-1}\right)v_0\right] + \eta v_0\right\},$$
$$\tag{1.51}$$

where $\bar{\rho} \equiv \rho + x^*_{-1} + \eta$. Since x^*_{n+1}, x^*_n and x^*_{n-1} are maximizers of the value functions v_{n+1}, v_n and v_{n-1}, (1.51) implies:

$$\bar{\rho}v_{n+1} = 1 - \lambda^{-n-1} - \omega^* G\left(x^*_{n+1}\right) + x^*_{n+1}\left[v_{n+2} - v_{n+1}\right] + x^*_{-1}\left[a^*_{-1}\left(v_{-1} + \zeta\right) + \left(1 - a^*_{-1}\right)v_0\right] + \eta v_0,$$
$$\tag{1.52}$$

$$\bar{\rho}v_n \geq 1 - \lambda^{-n} - \omega^* G\left(x^*_{n+1}\right) + x^*_{n+1}\left[v_{n+1} - v_n\right] + x^*_{-1}\left[a^*_{-1}\left(v_{-1} + \zeta\right) + \left(1 - a^*_{-1}\right)v_0\right] + \eta v_0,$$

$$\bar{\rho}v_n \geq 1 - \lambda^{-n} - \omega^* G\left(x^*_{n-1}\right) + x^*_{n-1}\left[v_{n+1} - v_n\right] + x^*_{-1}\left[a^*_{-1}\left(v_{-1} + \zeta\right) + \left(1 - a^*_{-1}\right)v_0\right] + \eta v_0,$$

$$\bar{\rho}v_{n-1} = 1 - \lambda^{-n+1} - \omega^* G\left(x^*_{n-1}\right) + x^*_{n-1}\left[v_n - v_{n-1}\right] + x^*_{-1}\left[a^*_{-1}\left(v_{-1} + \zeta\right) + \left(1 - a^*_{-1}\right)v_0\right] + \eta v_0.$$

Now taking differences with $\bar{\rho}v_n$ and using the definitions of δ_ns, we obtain

$$\bar{\rho}\delta_{n+1} \leq \lambda^{-n}\left(1 - \lambda^{-1}\right) + x^*_{n+1}\left(\delta_{n+2} - \delta_{n+1}\right)$$
$$\bar{\rho}\delta_n \geq \lambda^{-n+1}\left(1 - \lambda^{-1}\right) + x^*_{n-1}\left(\delta_{n+1} - \delta_n\right).$$

Therefore,

$$\left(\bar{\rho} + x^*_{n-1}\right)\left(\delta_{n+1} - \delta_n\right) \leq -k_n + x^*_{n+1}\left(\delta_{n+2} - \delta_{n+1}\right), \tag{1.53}$$

where

$$k_n \equiv \left(\lambda - 1\right)^2 \lambda^{-n-1} > 0.$$

Now to obtain a contradiction, suppose that $\delta_{n+1} - \delta_n \geq 0$. From (1.53), this implies $\delta_{n+2} - \delta_{n+1} > 0$ since k_n is strictly positive. Repeating this argument successively, we have that if $\delta_{n'+1} - \delta_{n'} \geq 0$, then $\delta_{n+1} - \delta_n > 0$ for all $n \geq n'$. However, we know from Part 2 of the proposition that $\{v_n\}_{n=0}^{\infty}$ is strictly increasing and converges to a constant v_∞. This implies that $\delta_n \downarrow 0$, which contradicts the hypothesis that $\delta_{n+1} - \delta_n \geq 0$ for all $n \geq n' \geq 0$, and establishes that $x^*_{n+1} \leq x^*_n$. To see that the inequality is strict when $x^*_n > 0$, it suffices to note that we have already established (1.50), i.e., $\delta_{n+1} - \delta_n < 0$, thus if equation (1.28) has a positive solution, then we necessarily have $x^*_{n+1} < x^*_n$.

59

We next prove that $x_0^* \geq x_{-1}^*$ and then show that under the additional condition $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$, this inequality is strict.

Proof of $x_0^* \geq x_{-1}^*$: Suppose first that $\zeta > v_1 - v_0$. Then (1.27) implies $a_{-1}^* = 0$, and (1.25) can be written as

$$\rho v_0 = -\omega^* G\left(x_0^*\right) + x_0^*\left[v_{-1} + v_1 - 2v_0\right]. \tag{1.54}$$

We have $v_0 \geq 0$ from Part 2 of the proposition. Suppose $v_0 > 0$. Then (1.54) implies $x_0^* > 0$ and

$$
\begin{aligned}
v_{-1} + v_1 - 2v_0 &> 0 \tag{1.55}\\
v_1 - v_0 &> v_0 - v_{-1}.
\end{aligned}
$$

This inequality combined with $a_{-1}^* = 0$, (1.30) and (1.37) yields $x_0^* > x_{-1}^*$. Suppose next that $v_0 = 0$. Inequality (1.55) now holds as a weak inequality and implies that $x_0^* \geq x_{-1}^*$. Moreover, since $G\left(\cdot\right)$ is strictly convex and x_0^* is given by (1.30), (1.54) then implies $x_0^* = 0$ and thus $x_{-1}^* = 0$.

We next show that when $\zeta \leq v_1 - v_0$, $x_0^* \geq x_{-1}^*$. In this case, $a_{-1}^* = 1$ is an optimal policy, so that

$$
\begin{aligned}
\rho v_0 &= -\omega^* G\left(x_0^*\right) + x_0^*\left[v_1 - v_0\right] + x_0^*\left[v_{-1} - v_0\right]\\
\rho v_{-1} &\geq -\omega^* G\left(x_0^*\right) + x_0^*\left[v_1 - v_{-1} - \zeta\right] + \eta\left[v_0 - v_{-1}\right].
\end{aligned}
$$

Subtracting the second expression from the first, we obtain

$$\rho\left[v_0 - v_{-1}\right] \leq x_0^*\left[v_{-1} + \zeta - v_0\right] + \left(x_0^* + \eta\right)\left[v_{-1} - v_0\right],$$

and therefore

$$\left[v_0 - v_{-1}\right] \leq \left[v_{-1} + \zeta - v_0\right].$$

Part 2 of the proposition implies that $v_{-1} \leq v_0$, and therefore $v_{-1} + \zeta \geq v_0$. Next observe that with $a_{-1}^* = 1$, (1.30) and (1.37) imply that $x_0^* \geq x_{-1}^*$ if and only if $v_1 - v_0 \geq v_1 - v_{-1} - \zeta$, or equivalently if and only if $v_{-1} + \zeta \geq v_0$. Thus we have established that $x_0^* \geq x_{-1}^*$ both when $\zeta > v_1 - v_0$ and when $\zeta \leq v_1 - v_0$. $\square$

We now have the following intermediate lemma.

Lemma 1.2 *Suppose that $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$, then $x_0^* > 0$ and $v_0 > 0$.*

Proof. Suppose, to obtain a contradiction, that $x_0^* = 0$. The first part of the proof then implies that $x_{-1}^* = 0$. Then (1.24) implies

$$\rho v_1 \geq 1 - \lambda + \eta\left[v_0 - v_1\right].$$

Equation (1.25) together with $x_0^* = 0$ gives $v_0 = 0$, and hence

$$v_1 - v_0 \geq \frac{1 - \lambda^{-1}}{\rho + \eta}.$$

Combined with this inequality, (1.30) implies

$$
\begin{aligned}
x_0^* &\geq \max\left\{G'^{-1}\left(\frac{1 - \lambda^{-1}}{\omega^*\left(\rho + \eta\right)}\right), 0\right\},\\
&\geq \max\left\{G'^{-1}\left(\frac{1 - \lambda^{-1}}{\rho + \eta}\right), 0\right\},
\end{aligned}
$$

where the second inequality follows from the fact that $\omega^* \leq 1$. The assumption that $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$ then implies $x_0^* > 0$, thus leading to a contradiction and establishing that $x_0^* > 0$. Strict convexity of $G\left(\cdot\right)$ together with $x_0^* > 0$ then implies $v_0 > 0$. ■

Proof of $x_0^* > x_{-1}^*$ when $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$ and $\zeta > 0$: Given Lemma 1.2, $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$ implies that $x_0^* > 0$. Then the first part of the proof implies that when $\zeta > v_1 - v_0$, $x_0^* > x_{-1}^*$. Next suppose that $0 < \zeta < v_1 - v_0$. Then the same argument as above implies that $x_0^* > x_{-1}^*$ if and only if $v_1 - v_0 > v_1 - v_{-1} - \zeta$, or equivalently if and only if $v_{-1} + \zeta > v_0$. Suppose this is not the case. Then from the first part of the proof, we have that $x_0^* = x_{-1}^* = 0$, and thus $v_{-1} = v_0 = 0$, which implies $v_{-1} + \zeta > v_0$ and thus $x_0^* > x_{-1}^*$. This yields a contradiction and completes the proof that $x_0^* > x_{-1}^*$ when $\zeta > 0$ and $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$. $\square$

Proof of $x_0^* > x_1^*$: To prove that $x_0^* > x_1^*$, let us write the value functions v_2, v_1 and v_0 as in (1.52):

$$\bar{\rho}v_2 = 1 - \lambda^{-2} - \omega^* G\left(x_2^*\right) + x_2^*\left[v_3 - v_2\right] + x_{-1}^*\left[a_{-1}^*\left(v_{-1} + \zeta\right) + \left(1 - a_{-1}^*\right)v_0\right] + \eta v_0,$$

$$\bar{\rho}v_1 \geq 1 - \lambda^{-1} - \omega^* G\left(x_2^*\right) + x_2^*\left[v_2 - v_1\right] + x_{-1}^*\left[a_{-1}^*\left(v_{-1} + \zeta\right) + \left(1 - a_{-1}^*\right)v_0\right] + \eta v_0,$$

$$\bar{\rho}v_1 \geq 1 - \lambda^{-1} - \omega^* G\left(x_0^*\right) + x_0^*\left[v_2 - v_1\right] + x_{-1}^*\left[a_{-1}^*\left(v_{-1} + \zeta\right) + \left(1 - a_{-1}^*\right)v_0\right] + \eta v_0,$$

$$\bar{\rho}v_0 = -\omega^* G\left(x_0\right) + x_0^*\left[v_1 - v_0\right] + \eta v_0 + x_{-1}^* v_0 + x_0^*\left[v_{-1} - v_0\right].$$

Now taking differences with $\bar{\rho}v_n$ and using the definitions of δ_ns as in (1.50), we obtain

$$\bar{\rho}\delta_2 \leq \lambda^{-1}\left(1 - \lambda^{-1}\right) + x_2^*\left(\delta_3 - \delta_2\right), \tag{1.56}$$

$$\bar{\rho}\delta_1 \geq \left(1 - \lambda^{-1}\right) + x_0^*\left(\delta_2 - \delta_1\right) + x_{-1}^*\left[a_{-1}^*\left(v_{-1} + \zeta\right) + \left(1 - a_{-1}^*\right)v_0 - v_0\right] - x_0^*\left[v_{-1} - v_0\right],$$

$$\bar{\rho}\delta_1 \geq \left(1 - \lambda^{-1}\right) + x_0^*\left(\delta_2 - \delta_1\right) + x_{-1}^* a_{-1}^* \zeta + \left(x_{-1}^* a_{-1}^* - x_0^*\right)\left[v_{-1} - v_0\right],$$

$$\bar{\rho}\delta_1 \geq \left(1 - \lambda^{-1}\right) + x_0^*\left(\delta_2 - \delta_1\right) + \left(x_{-1}^* a_{-1}^* - x_0^*\right)\left[v_{-1} - v_0\right].$$

Next recall from Part 2 that $v_{-1} - v_0 \leq 0$. Moreover, the first part of the first part of the proof has established that $x_{-1}^* - x_0^* \leq 0$. Combining this with $a_{-1}^* \leq 1$ establishes that $\left[x_{-1}^* - x_0\right]\left[v_{-1} - v_0\right] \geq 0$, and the last inequality then implies

$$\bar{\rho}\delta_1 \geq \left(1 - \lambda^{-1}\right) + x_0^*\left(\delta_2 - \delta_1\right).$$

Now combining this inequality with the first inequality of (1.56), we obtain

$$\left(\bar{\rho} + x_0^*\right)\left(\delta_2 - \delta_1\right) \leq -\left(1 - \lambda^{-1}\right)^2 + x_2^*\left(\delta_3 - \delta_2\right). \tag{1.57}$$

Part 2 has already established $\delta_2 > \delta_3$, so that the right-hand side is strictly negative, therefore, we must have $\delta_2 - \delta_1 < 0$, which implies that $x_0^* > x_1^*$ and completes the proof. $\square$

The above results together complete the proof of Part 3. ∎

Part 4: Uniqueness of the Invariant Distribution.

Lemma 1.3 *Consider a uniform IPR policy $\left\langle \eta^{uni}, \zeta^{uni}\right\rangle$ and a corresponding steady-state equilibrium $\left\langle \mu^*, \mathbf{v}, a_{-1}^*, \mathbf{x}^*, \omega^*, g^*\right\rangle$. Then, there exists $n^* \in \mathbb{N}$ such that $x_n^* = 0$ for all $n \geq n^*$.*

Proof. The first-order condition of the maximization of the value function (1.24) implies:

$$G'\left(x_n\right) \geq \frac{v_{n+1} - v_n}{\omega^*} \text{ and } x_n \geq 0,$$

with complementary slackness. $G'(0)$ is strictly positive by assumption. If $\left(v_{n+1} - v_n\right)/\omega^* < G'(0)$, then $x_n = 0$. The second part of the proposition implies that $\{v_n\}_{n=-1}^{\infty}$ is a convergent and thus a Cauchy sequence, which implies that there exists $\exists n^* \in \mathbb{N}$ such that $v_{n+1} - v_n < \omega^* G'(0)$ for all $n \geq n^*$. ∎

An immediate consequence of Lemma 1.3, combined with (1.31) is that $\mu_n = 0$ for all $n \geq n^*$ (since there is no innovation in industries with technology gap greater than n^*). Thus the law of motion of an industry can be represented by a finite Markov chain. Moreover, because after an innovation by a follower, all industries jump to the neck-and-neck state (when $a_{-1}^* = 0$) or to the technology gap of one (when $a_{-1}^* = 1$), this Markov chain is irreducible (and aperiodic), thus converges to a unique steady-state distribution of industries. More formally, there exists n^* such that $x_{n^*}^* = 0$ and $x_n^* = 0$ for all $n > n^*$. Combined with the fact $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$ and that either $\eta > 0$ or $x_{-1}^* > 0$, this implies that the states $n > n^*$ are transient and can be ignored. Consequently, $\{\mu_n^*\}_{n=0}^{\infty}$ forms a finite and irreducible Markov chain over the states $n = 0, 1, ..., n^*$. To see this, let $n^* = \min_{n \in \{0, .. ,n^{**}\}} \{n \in \mathbb{N} : v_{n+1} - v_n \leq \omega^* G'(0)\}$. Such an n^* exists, since the set $\{0, ..., n^{**}\}$ is finite and nonempty because of the assumption that $G'^{-1}\left(\left(1 - \lambda^{-1}\right)/\left(\rho + \eta\right)\right) > 0$. Then by construction $x_n^* > 0$ for all $n < n^*$ and $x_{n^*}^* = 0$ as desired. Now denoting the probability of being in state $\tilde{n}$ starting in state n after τ periods by $P^{\tau}(n, \tilde{n})$, we have that $\lim_{\tau \to \infty} P^{\tau}(n, \tilde{n}) = 0$ for all $\tilde{n} > n^*$ and for all n. Thus we can focus on the finite Markov chain over the states $n = 0, 1, ..., n^*$, and $\{\mu_n^*\}_{n=0}^{n^*}$ is the limiting (invariant) distribution of this Markov chain. Given a_{-1}^* and $\{x_n^*\}_{n=-1}^{n^*}$, $\{\mu_n^*\}_{n=0}^{n^*}$ is uniquely defined. Moreover, the underlying Markov chain is irreducible (since $x_n^* > 0$ for $n = 0, 1, ..., n^* - 1$, so that all states communicate with $n = 0$ or $n = 1$). Therefore, by Theorem 11.2 in Stokey, Lucas and Prescott (1989, p. 62) there exists a unique stationary distribution $\{\mu_n^*\}_{n=0}^{\infty}$. ∎

Proof of Proposition 1.3

We prove this proposition using two crucial lemmas.

Lemma 1.4 *Consider the state-dependent IPR policy $\langle \boldsymbol{\eta}, \boldsymbol{\zeta} \rangle$, and suppose that $\langle \boldsymbol{\mu}^*, \mathbf{v}, a_{-1}^*, \mathbf{x}^*, \omega^*, g^* \rangle$ is a steady-state equilibrium. Then there exists a state $n^* \in \mathbb{N}$ such that $\mu_n^* = 0$ for all $n \geq n^*$.*

Proof. There are two cases to consider. First, suppose that $\{v_n\}_{n \in \mathbb{Z}_+}$ is strictly increasing. Then it follows from the proof of Lemma 1.3 that there exists a state $n^* \in \mathbb{N}$ such that $x_n^* = 0$ for all $n \geq n^*$, and as in the proof of Part 4 of Proposition 1.2, states $n \geq n^*$ are transient (i.e., $\lim_{\tau \to \infty} P^{\tau}(n, \tilde{n}) = 0$ for all $\tilde{n} > n^*$ and for all n), so $\mu_n^* = 0$ for all $n \geq n^*$.

Second, in contrast to the first case, suppose that there exists some $n^{**} \in \mathbb{Z}_+$ such that $v_{n^{**}} \geq v_{n^{**}+1}$. Then, let $n^* = \min_{n \in \{0, ,n^{**}\}} \{n \in \mathbb{N} : v_{n+1} - v_n \leq \omega^* G'(0)\}$, which is again well defined. Then, optimal R&D decision (1.28) immediately implies that $x_n^* > 0$ for all states with $n < n^*$, and since $x_{n^*}^* = 0$, all states $n > n^*$ are transient and $\lim_{\tau \to \infty} P^{\tau}(n, \tilde{n}) = 0$ for all $\tilde{n} > n^*$ and for all n, completing the proof. ∎

Lemma 1.5 *Consider the state-dependent IPR policy $\langle \eta, \zeta \rangle$ and suppose that the labor share and the R&D policies of all other firms are given by $\mathbf{z} = \langle \tilde{\omega}, \tilde{\mathbf{a}}, \tilde{\mathbf{x}} \rangle$. Then the dynamic optimization problem of an individual firm leads to a unique value function $\mathbf{v}[\mathbf{z}] : \mathbb{Z} \to \mathbb{R}_+$ and optimal R&D policies $\hat{\mathbf{A}}[\mathbf{z}] : \mathbb{Z}_- \setminus \{0\} \rightrightarrows [0,1]$ and $\hat{\mathbf{X}}[\mathbf{z}] : \mathbb{Z} \rightrightarrows [0, \bar{x}]$ are compact and convex-valued for each $\mathbf{z} \in \mathbf{Z}$ and upper hemi-continuous in $\mathbf{z}$ (where $\mathbf{v}[\mathbf{z}] \equiv \{v_n[\mathbf{z}]\}_{n=-1}^{\infty}$, $\hat{\mathbf{A}}[\mathbf{z}] \equiv \left\{\hat{A}_n[\mathbf{z}]\right\}_{n=-\infty}^{-1}$ and $\hat{\mathbf{X}}[\mathbf{z}] \equiv \left\{\hat{X}_n[\mathbf{z}]\right\}_{n=-1}^{\infty}$).*

Proof. The proof follows closely that of Lemma 1.1. In particular, again using uniformization, the maximization problem of an individual firm can be written as a contraction mapping similar to (1.44) there. The finiteness of the transition probabilities follows, since $\psi_n\left(\boldsymbol{\xi} \mid \tilde{\boldsymbol{\xi}}\right) \leq \psi \equiv 2\bar{x} + \max_n \{\eta_n\} < \infty$ (this is a consequence of the fact that $\bar{x}$ defined in (1.11) is finite and $\max_n \{\eta_n\}$ is finite, since each $\eta_n \in \mathbb{R}_+$ and by assumption, there exists $\bar{n} < \infty$ such that $\eta_n = \eta_{\bar{n}}$). This contraction mapping uniquely determines the value function $\mathbf{v}[\mathbf{z}] : \mathbb{Z} \to \mathbb{R}_+$.

Berge's Maximum Theorem (Aliprantis and Border, 1999, Theorem 16.31, p. 539) again implies that each of $\hat{A}_n(\mathbf{z})$ for $n \in \mathbb{Z}_- \setminus \{0\}$ and $\hat{X}_n(\mathbf{z})$ for $n \in \mathbb{Z}$ is upper hemi-continuous in $\mathbf{z} = \langle \tilde{\omega}, \tilde{\mathbf{a}}, \tilde{\mathbf{x}} \rangle$, and

moreover, since v_n for $n \in \mathbb{Z}$ is concave in a_n and x_n, the maximizers of $\mathbf{v}[\mathbf{z}]$, $\hat{\boldsymbol{A}} \equiv \left\{ \hat{A}_{-n} \right\}_{n=1}^{\infty}$ and $\hat{\mathbf{X}} \equiv \left\{ \hat{X}_n \right\}_{n=-\infty}^{\infty}$, are nonempty, compact and convex-valued. ∎

Now using the previous two lemmas, we can establish the existence of a steady-state equilibrium. This part of the proof follows that of Proposition 1.2 closely. Fix $\mathbf{z} = \left\langle \tilde{\omega}, \{\tilde{a}_n\}_{n=-\infty}^{-1}, \{\tilde{x}_n\}_{n=-\infty}^{\infty} \right\rangle$, and define $\mathbf{Z} \equiv [0,1] \times \prod_{n=-1}^{\infty} [0,1] \times \prod_{n=-\infty}^{\infty} [0,\bar{x}]$. Again by Tychonoff's Theorem, $\mathbf{Z}$ is compact in the product topology. Then consider the mapping $\Phi \colon \mathbf{Z} \rightrightarrows \mathbf{Z}$ constructed as $\Phi \equiv \left(\varphi, \hat{\mathbf{A}}, \hat{\mathbf{X}} \right)$, where φ is given by (1.45) and $\hat{\mathbf{A}}$ and $\hat{\mathbf{X}}$ are defined in Lemma 1.5. Clearly Φ maps $\mathbf{Z}$ into itself. Moreover, as in the proof of Proposition 1.2, $\mathbf{Z}$ is nonempty, convex, and a subset of a locally convex Hausdorff space. The proof of Lemma 1.5 then implies that Φ has closed graph in the product topology and is nonempty, compact and convex-valued in $\mathbf{z}$. Consequently, the Kakutani-Fan-Glicksberg Fixed Point Theorem again applies and implies that Φ has a fixed point $\mathbf{z}^* \in \Phi(\mathbf{z}^*)$. The argument that the fixed point $\mathbf{z}^*$ corresponds to a steady-state equilibrium is identical to that in Proposition 1.2, and follows from the fact that within argument identical to that of Lemma 1.2, $G'^{-1}\left(\left(1 - \lambda^{-1}\right) / (\rho + \eta_1) \right) > 0$ implies that $x_0^* > 0$. The result that $\omega^* < 1$ then follows immediately. Finally, as in the proof of Proposition 1.2, either $\eta_1 > 0$ or $x_{-1}^* > 0$ is sufficient for $g^* > 0$. ∎

References

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith and Peter Howitt (2005) "Competition and Innovation: an Inverted-U Relationship" *Quarterly Journal of Economics*, 120, pp. 701-728.

Aghion, Philippe, Christopher Harris, Peter Howitt and John Vickers (2001) "Competition, Imitation and Growth with Step-by-Step Innovation" *Review of Economic Studies*, 68, pp. 467-492 .

Aghion, Philippe, Christopher Harris and John Vickers (1997) "Step-by-step Competition, Innovation and Growth: An Example" *European Economic Review*, 41, pp. 771-782.

Aghion, Philippe and Peter Howitt (1992) "A Model of Growth Through Creative Destruction" *Econometrica*, 110, 323-351.

Aliprantis, Charalambos D. and Kim C. Border (1999) *Infinite Dimensional Analysis: Hitchhiker's Guide*, Springer-Verlag, New York.

Arrow, Kenneth (1962) "Economic Welfare and the Allocation of Resources for Invention" in R.R. Nelson, ed. *Universities-National Bureau of Economic Research Conference Series; The Rate and Direction of Economic Activities: Economic and Social Factors*, Princeton University Press, New York.

Baye, Michael R., and Heidrun C. Hoppe (2003) "The strategic equivalence of rent-seeking, innovation, and patent-race games," *Games and Economic Behavior*, 44, 217-226.

Bloom, Nick, Mark Schankerman and John Van Reenen (2005) "Identifying Technology Spillovers and Product Rivalry," CEP Discussion Paper, No. 675.

Boldrin, Michele and David K. Levine (2001) "Perfectly Competitive Innovation," Univesity of Minnesota and UCLA mimeo.

Boldrin, Michele and David K. Levine (2004) "Innovation and the Size of the Market," Univesity of Minnesota and UCLA mimeo.

Budd, Christopher, Christopher Harris and John Vickers (1993) "A Model of the Evolution of Duopoly: Does the Asymmetry Between Firms Tend to Increase or Decrease?" *Review of Economic Studies*, 60, 540-573.

Bureau of Economic Analysis (2004) "Annual Revision of the National Income and Product Accounts" *Bureau of Economic Analysis News*. BEA 04-31. www.bea.gov/bea/rels.htm

Choi, Jay P. (1991) "Dynamic R&D Competition under Hazard Rate Uncertainty," *Rand Journal of Economics*, 22, 596-610.

Fudenberg, Drew, Richard Gilbert, Joseph Stiglitz and Jean Tirole (1983) "Preemption, leapfrogging and competition in patent races," *European Economic Review* 22, 3-31.

Fullerton, Robert and Preston McAfee (1999) "Auctioning entry into tournaments," *Journal of Political Economy*, 107, 573-605.

Gallini, Nancy T. (1992) "Patent Policy and Costly Imitation" *RAND Journal of Economics*, 23, 52-63.

Gallini, Nancy T. and Suzanne Scotchmer (2002) "Intellectual Property; When is it the Best Incentive Mechanism?" *Innovation Policy and the Economy*, 2, 51-78.

Gilbert, Richard. and Carl Shapiro (1990) "Optimal Patent Length and Breadth," *Journal of Industrial Economics*, 21(1), 106-112.

Green, Jerry and Suzanne Scotchmer (1995) "On the Division of Profit in Sequential Innovation," *Journal of Industrial Economics*, 26(1), 20-33.

Grossman, Gene and Elhanan Helpman (1990) *Innovation and Growth in the Global Economy*, Cambridge, MA, MIT Press.

Hall, Bronwyn H., Jerry Hausman and Zvi Griliches (1988) "Patents and R&D: Is There a Lag?" *International Economic Review*, 27, pp. 265-83.

Harris, Christopher and John Vickers (1985) "Perfect Equilibrium in a Model of Race" *Review of Economic Studies*, 52, 193-209.

Harris, Christopher and John Vickers (1987) "Racing with Uncertainty" *Review of Economic Studies*, 54, 1-21.

Hopenhayn, Hugo and Matthew Mitchell (2001) "Innovation, Variety and Patent Breadth" *RAND Journal of Economics*, 32, 152-166.

Hugo Hopenhayn, Llobet, Gerard and Matthew Mitchell (2006) "Rewarding Sequential Innovators: Prizes, Patents and Buyouts," *Journal of Political Economy*, 114, 1041-1068.

Judd, Kenneth (1999) *Numerical Methods in Economics*, MIT Press, Cambridge.

Kelley, John L. (1955) *General Topology*, Springer-Verlag, New York.

Klemperer, Paul (1990) "How Broad Should the Scope of Patent Protection Be?" *Journal*

of Industrial Economics, 21, 113-130.

Kortum, Samuel (1993). "Equilibrium R&D and the Patent-R&D Ratio: US Evidence" *American Economic Review,* 83, pp. 450-457.

Moscarini, Giuseppe and Francesco Squintani (2004) "Competitive Experimentation with Private Information," Cowles Foundation Discussion Paper, No. 1489.

National Science Foundation (2006) "Industrial R&D Employment in the United States and in U.S. Multinational Corporations" *Division of Science Resource Statistics.* NSF 05-302. Arlington, VA: National Science Foundation. www.bsf.gov/sbe/srs/stats.htm

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

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

Pakes, Ariel and Zvi Griliches (1980) "Patents and R&D at the Firm Level: A First Look." *Economic Letters,* 5, pp. 377-381.

Quah, Danny (2004) "24/7 Competitive Innovation," LSE Mimeo.

Reinganum, Jennifer (1981) "Dynamic Games of Innovation," *Journal of Economic Theory,* 25, 21-24.

Reinganum, Jennifer (1985) "Innovation and Industry Evolution," *Quarterly Journal of Economics,* 100, 81-100.

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

Ross, Sheldon (1996) *Stochastic Processes,* John Wiley & Sons, Inc.

Scotchmer, Suzanne (1999) "On the Optimality of the Patent System," *The Rand Journal of Economics,* 30, 181-196.

Scotchmer, Suzanne (2005) *Innovations and Incentives,* MIT Press, Cambridge, Massachusetts.

Segal, Ilya and Michael Whinston (2005) "Anti-Trust in Innovative Industries," National Bureau of Economic Research Working Paper.

Taylor, Curtis (1995) "Digging for golden carrots: an analysis of research tournaments,"

American Economic Review, 85, 872-890.

Tirole, Jean (1988) *The Theory of Industrial Organization*, MIT Press, Cambridge, Massachusetts.

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

Stokey, Nancy and Robert Lucas with E. Prescott (1989) *Recursive Methods in Economic Dynamics*, Harvard University Press.

Table 1. Benchmark Results

	Full IPR Protection without licensing	Full IPR Protection with licensing	Optimal Uniform IPR with licensing	Optimal State Dependent without licensing	Optimal State Dependent with licensing
λ	1.05	1.05	1.05	1.05	1.05
γ	0.35	0.35	0.35	0.35	0.35
η_1	0	0	0	0.71	0
η_2	0	0	0	0.08	0
η_3	0	0	0	0	0
η_4	0	0	0	0	0
η_5	0	0	0	0	0
ζ_1	∞	3.53	0	∞	0
ζ_2	∞	3.53	0	∞	0.98
ζ_3	∞	3.53	0	∞	1.93
ζ_4	∞	3.53	0	∞	1.97
ζ_5	∞	3.53	0	∞	1.98
$v_1 - v_0$	2.71	3.53	1.71	1.52	1.98
x_{-1}^*	0.22	0.25	0.27	0.12	0.32
x_0^*	0.35	0.41	0.27	0.25	0.30
μ_0^*	0.24	0	0	0.46	0
μ_1^*	0.33	0.48	0.51	0.19	0.44
μ_2^*	0.20	0.25	0.25	0.13	0.24
ω^*	0.95	0.93	0.94	0.96	0.94
Researcher ratio	0.032	0.026	0.027	0.028	0.039
$\ln C(0)$	33.78	35.47	35.57	34.20	36.16
g^*	0.0186	0.0258	0.0263	0.0204	0.0296
Welfare	683.0	719.8	722.0	692.1	735.1

Note: This table gives the results of the benchmark numerical computations with $\rho = 0.05$, $\lambda = 1.05$, $\gamma = 0.35$ under five different IPR policy regimes. It reports the steady-state equilibrium values of the difference in the values $v_1 - v_0$; the (annual) R&D rate of a follower that is one step behind, x_{-1}^*; the (annual) R&D rate of neck-and-neck competitors, x_0^*; fraction of industries in neck-and-neck competition, μ_0^*; fraction of industries at a technology gap of $n = 1, 2$; the value of "labor share," ω^*; the ratio of the labor force working in research; initial (annual) consumption, $C(0)$; the annual growth rate, g^*; and the welfare level according to equation (1.38). It also reports the welfare-maximizing uniform and state-dependent IPR policies with or without licensing. See text for details.

Table 2. $\lambda = 1.01$

	Full IPR Protection without licensing	Full IPR Protection with licensing	Optimal Uniform IPR with licensing	Optimal State Dependent without licensing	Optimal State Dependent with licensing
λ	1.01	1.01	1.01	1.01	1.01
γ	0.35	0.35	0.35	0.35	0.35
η_1	0	0	0	3.14	0.06
η_2	0	0	0	0.23	0
η_3	0	0	0	0	0
η_4	0	0	0	0	0
η_5	0	0	0	0	0
ζ_1	∞	0.19	0	∞	0.04
ζ_2	∞	0.19	0	∞	0.10
ζ_3	∞	0.19	0	∞	0.10
ζ_4	∞	0.19	0	∞	0.10
ζ_5	∞	0.19	0	∞	0.10
$v_1 - v_0$	0.14	0.19	0.09	0.08	0.10
x^*_{-1}	1.08	1.27	1.29	0.64	1.51
x^*_0	1.67	1.95	1.29	1.25	1.38
μ^*_0	0.25	0	0	0.45	0.01
μ^*_1	0.33	0.50	0.50	0.20	0.45
μ^*_2	0.19	0.25	0.25	0.14	0.23
ω^*	0.99	0.99	0.99	0.99	0.99
Researcher ratio	0.008	0.007	0.007	0.007	0.010
$\ln C(0)$	10.28	11.88	11.90	10.66	12.66
g^*	0.0186	0.0257	0.0258	0.0203	0.0293
Welfare	213.1	247.8	248.3	221.4	265.0

Note: This table gives the results of the numerical computations with $\rho = 0.05$, $\lambda = 1.01$, $\gamma = 0.35$ under five different IPR policy regimes. It reports the steady-state equilibrium values of the difference in the values $v_1 - v_0$; the (annual) R&D rate of a follower that is one step behind, x^*_{-1}; the (annual) R&D rate of neck-and-neck competitors, x^*_0; fraction of industries in neck-and-neck competition, μ^*_0; fraction of industries at a technology gap of $n = 1, 2$; the value of "labor share," ω^*; the ratio of the labor force working in research; initial (annual) consumption, $C(0)$; the annual growth rate, g^*; and the welfare level according to equation (1.38). It also reports the welfare-maximizing uniform and state-dependent IPR policies with or without licensing. See text for details.

Table 3. $\lambda = 1.2$

	Full IPR Protection without licensing	Full IPR Protection with licensing	Optimal Uniform IPR with licensing	Optimal State Dependent without licensing	Optimal State Dependent with licensing
λ	1.20	1.20	1.20	1.20	1.20
γ	0.35	0.35	0.35	0.35	0.35
η_1	0	0	0	0.19	0
η_2	0	0	0	0.08	0
η_3	0	0	0	0.05	0
η_4	0	0	0	0.05	0
η_5	0	0	0	0.05	0
ζ_1	∞	25.07	0	∞	0
ζ_2	∞	25.07	0	∞	5.52
ζ_3	∞	25.07	0	∞	10.95
ζ_4	∞	25.07	0	∞	13.64
ζ_5	∞	25.07	0	∞	14.98
$v_1 - v_0$	20.29	25.07	13.86	9.62	14.98
x_{-1}^*	0.06	0.068	0.08	0.02	0.09
x_0^*	0.10	0.12	0.08	0.06	0.09
μ_0^*	0.22	0	0	0.56	0
μ_1^*	0.32	0.45	0.52	0.24	0.45
μ_2^*	0.20	0.26	0.26	0.10	0.24
ω^*	0.81	0.74	0.78	0.92	0.76
Researcher ratio	0.069	0.057	0.070	0.037	0.089
$\ln C(0)$	114.78	116.79	117.17	115.49	117.30
g^*	0.0186	0.0265	0.0282	0.0189	0.0306
Welfare	2303.0	2346.5	2354.7	2317.3	2358.3

Note: This table gives the results of the numerical computations with $\rho = 0.05$, $\lambda = 1.2$, $\gamma = 0.35$ under five different IPR policy regimes. It reports the steady-state equilibrium values of the difference in the values $v_1 - v_0$; the (annual) R&D rate of a follower that is one step behind, x_{-1}^*; the (annual) R&D rate of neck-and-neck competitors, x_0^*; fraction of industries in neck-and-neck competition, μ_0^*; fraction of industries at a technology gap of $n = 1, 2$; the value of "labor share," ω^*; the ratio of the labor force working in research; initial (annual) consumption, $C(0)$; the annual growth rate, g^*; and the welfare level according to equation (1.38). It also reports the welfare-maximizing uniform and state-dependent IPR policies with or without licensing. See text for details.

Table 4. $\gamma = 0.1$

	Full IPR Protection without licensing	Full IPR Protection with licensing	Optimal Uniform IPR with licensing	Optimal State Dependent without licensing	Optimal State Dependent with licensing
λ	1.05	1.05	1.05	1.05	1.05
γ	0.1	0.1	0.1	0.1	0.1
η_1	0	0	0	3.18	0
η_2	0	0	0	0.04	0
η_3	0	0	0	0	0
η_4	0	0	0	0	0
η_5	0	0	0	0	0
ζ_1	∞	3.12	0	∞	0
ζ_2	∞	3.12	0	∞	0.62
ζ_3	∞	3.12	0	∞	1.74
ζ_4	∞	3.12	0	∞	1.82
ζ_5	∞	3.12	0	∞	1.84
$v_1 - v_0$	2.21	3.12	1.64	0.49	1.88
x_{-1}^*	0.27	0.28	0.29	0.19	0.29
x_0^*	0.29	0.31	0.29	0.25	0.29
μ_0^*	0.31	0	0	0.77	0
μ_1^*	0.33	0.50	0.50	0.10	0.49
μ_2^*	0.17	0.25	0.25	0.06	0.25
ω^*	0.94	0.92	0.92	0.98	0.92
Researcher ratio	0.008	0.008	0.008	0.003	0.010
$\ln C(0)$	34.07	36.16	36.20	34.94	36.31
g^*	0.0186	0.0278	0.0280	0.0222	0.0286
Welfare	688.8	734.2	735.1	707.7	737.6

Note: This table gives the results of the numerical computations with $\rho = 0.05$, $\lambda = 1.05$, $\gamma = 0.1$ under five different IPR policy regimes. It reports the steady-state equilibrium values of the difference in the values $v_1 - v_0$; the (annual) R&D rate of a follower that is one step behind, x_{-1}^*; the (annual) R&D rate of neck-and-neck competitors, x_0^*; fraction of industries in neck-and-neck competition, μ_0^*; fraction of industries at a technology gap of $n = 1, 2$; the value of "labor share," ω^*; the ratio of the labor force working in research; initial (annual) consumption, $C(0)$; the annual growth rate, g^*; and the welfare level according to equation (1.38). It also reports the welfare-maximizing uniform and state-dependent IPR policies with or without licensing. See text for details.

71

	Full IPR Protection without licensing	Full IPR Protection with licensing	Optimal Uniform IPR with licensing	Optimal State Dependent without licensing	Optimal State Dependent with licensing
λ	1.05	1.05	1.05	1.05	1.05
γ	0.6	0.6	0.6	0.6	0.6
η_1	0	0	0	0.61	0.01
η_2	0	0	0	0.18	0
η_3	0	0	0	0.07	0
η_4	0	0	0	0.03	0
η_5	0	0	0	0	0
ζ_1	∞	5.26	5.26	∞	0
ζ_2	∞	5.26	5.26	∞	0.62
ζ_3	∞	5.26	5.26	∞	1.52
ζ_4	∞	5.26	5.26	∞	1.98
ζ_5	∞	5.26	5.26	∞	2.35
$v_1 - v_0$	4.40	5.26	5.26	2.47	2.35
x^*_{-1}	0.13	0.18	0.18	0.03	0.31
x^*_0	0.64	0.85	0.85	0.27	0.25
μ^*_0	0.09	0	0	0.29	0.01
μ^*_1	0.26	0.42	0.42	0.12	0.29
μ^*_2	0.19	0.25	0.25	0.09	0.15
ω^*	0.94	0.94	0.94	0.94	0.93
Researcher ratio	0.073	0.044	0.044	0.084	0.097
$\ln C(0)$	33.19	33.88	33.88	33.91	35.48
g^*	0.0186	0.0198	0.0198	0.0229	0.0303
Welfare	671.3	685.6	685.6	687.3	721.6

Note. This table gives the results of the numerical computations with $\rho = 0.05$, $\lambda = 1.05$, $\gamma = 0.6$ under five different IPR policy regimes. It reports the steady-state equilibrium values of the difference in the values $v_1 - v_0$; the (annual) R&D rate of a follower that is one step behind, x^*_{-1}; the (annual) R&D rate of neck-and-neck competitors, x^*_0; fraction of industries in neck-and-neck competition, μ^*_0; fraction of industries at a technology gap of $n = 1, 2$; the value of "labor share," ω^*; the ratio of the labor force working in research; initial (annual) consumption, $C(0)$; the annual growth rate, g^*; and the welfare level according to equation (1.38). It also reports the welfare-maximizing uniform and state-dependent IPR policies with or without licensing. See text for details.
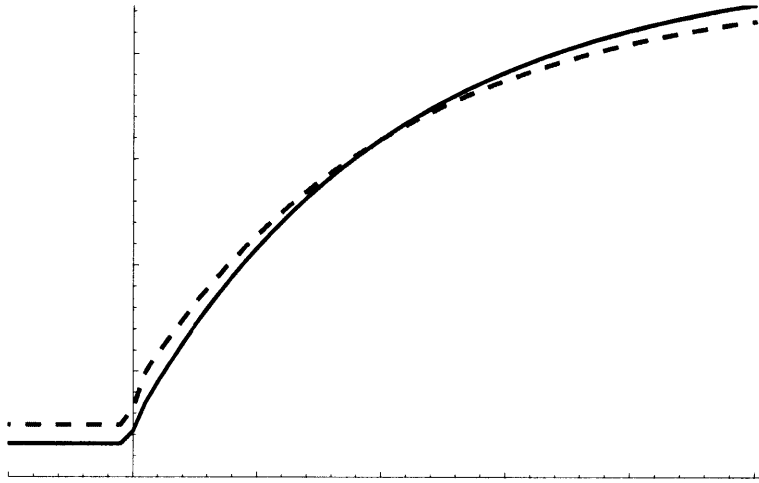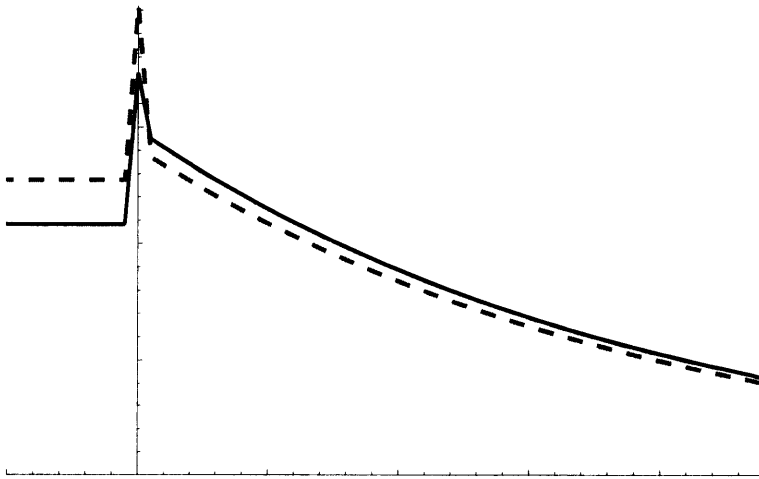
FIGURE 1. VALUE FUNCTIONS.
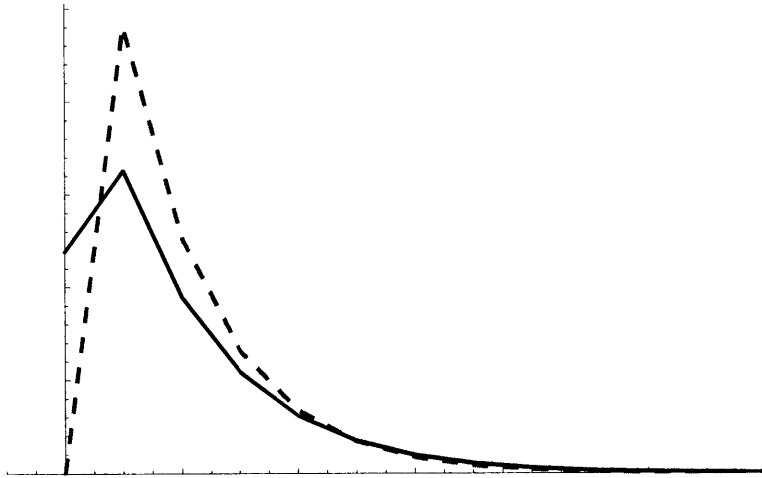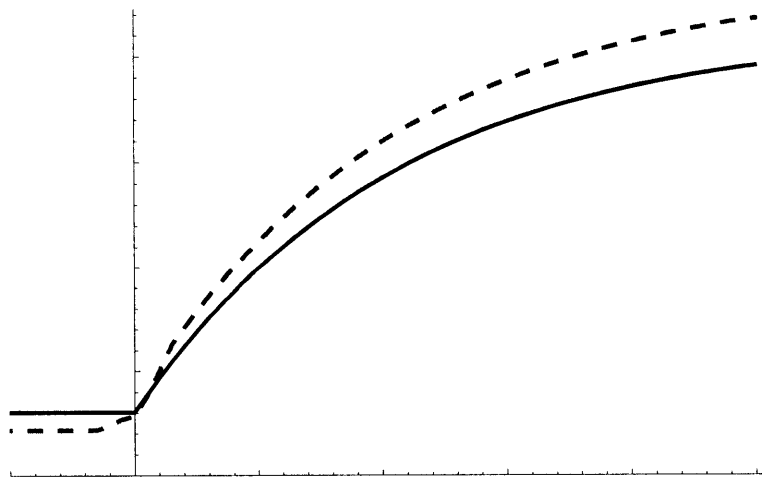

FIGURE 2. R&D EFFORTS.
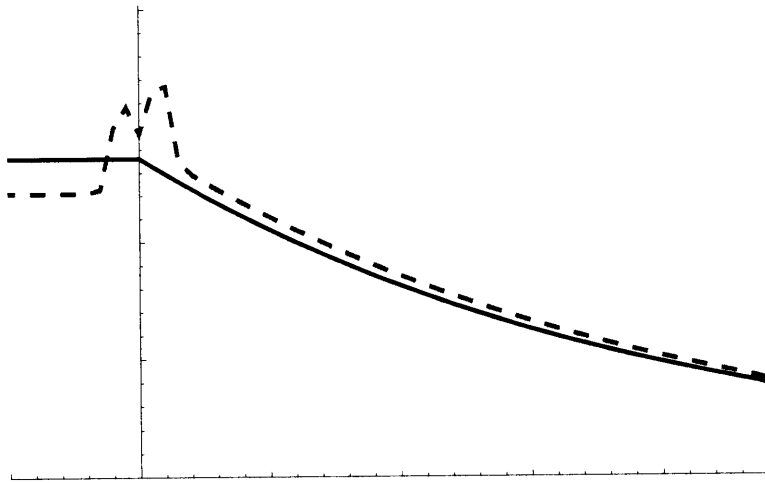
73

FIGURE 3. INDUSTRY SHARES.

FIGURE 4. VALUE FUNCTIONS.
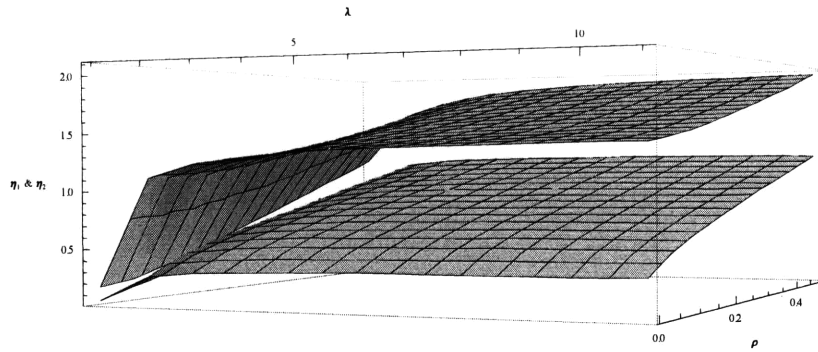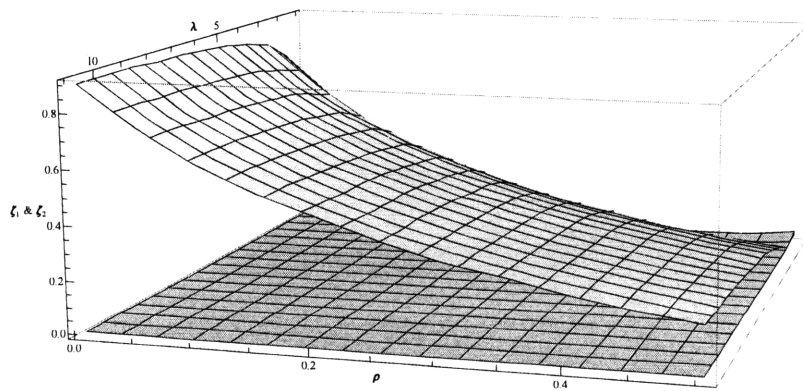
FIGURE 5. R&D EFFORTS.
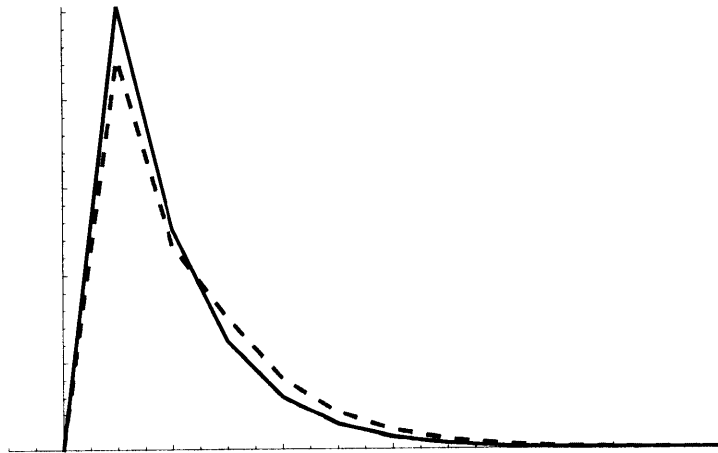
FIGURE 6. INDUSTRY SHARES.

η_1 −Upper Layer η_2 −Lower Layer

FIGURE 7. WELFARE MAXIMIZING FLOW RATES.

ζ_1 −Lower Layer ζ_2 −Upper Layer

FIGURE 8. WELFARE MAXIMIZING LICENSE FEES.

Chapter 2

Firm Size, Innovation Dynamics and Growth: A Reduced Form Analysis

2.1 Introduction

In his earlier work, Joseph Schumpeter (1934) claimed that the major source of innovation were small firms operated by wild-spirited entrepreneurs. However, in his later work, Schumpeter (1942) argued that the main innovators in an economy were actually the big firms that posess the required resources for important R&D laboratories and projects. His intriguing theories, also known as the Schumpeter Mark I and Mark II theories,[1] have raised many policy questions that are still open to debate. In particular, does firm size matter for innovation? Are big firms more R&D intensive as Schumpeter argued in his second theory? What are the implications of these facts for firm growth? These questions are at the heart of any policy debate which targets technological development through innovation. This chapter contributes to this debate by uncovering some important reduced form facts using Compustat and USPTO patent data. These facts would potentially form the basis for theoretical studies on the subject.

The first reduced form relationship considered here is the relationship between firm size and firm growth. The previous literature on firm growth is mainly centered around Gibrat's (1931) claim that the growth rate of a firm is independent of its size. I repeat a similar exercise with the Compustat dataset in order to generate an independent, but comparable measure of the relationship between firm size and growth. Doing so, I also address certain caveats that the previous empirical literature has overlooked. My results tend to corroborate the findings of

---

[1] Nelson and Winter (1982a,b) Kamien and Schwartz (1975, 1982)

previous studies which suggested that smaller firms grow faster on average. This effect becomes even sharper once unobserved firm heterogeneity is controlled for. These results are robust to, among other things, the survival bias and the selection of small firms in the Compustat sample.

The second reduced form relationship investigated is that between firm size and R&D intensity, defined as R&D spending over sales. In his seminal book, Joseph Schumpeter (1942) claimed that firm size was crucial for both R&D intensity and innovation suggesting that large firms have a size advantage. The early subsequent empirical studies had mostly supportive conclusions to Schumpeter's theory. However, more recent studies from the late 80s and early 90s have utilized larger datasets and addressed several econometric concerns, and have generally not found a systematic relationship between R&D intensity and firm size. My results on this question differ significantly from the previous literature. I show that the R&D intensity decreases significantly with firm size, implying that R&D spending increases less than proportional with firm size (as measured by sales). This result is robust to, among numerous other specifications, the problem of missing R&D entries in Compustat and the sample selection problem of small firms.

The last but most innovative reduced form analysis addresses the relationship between firm size and innovation quality. Empirical studies that have focused only on innovation frequency and neglected innovation quality, have been inadequate to capture the true effect of innovations on technological progress. One single, major innovation could be much more important for technological progress than many incremental innovations. The main empirical challenge lies in measuring innovation quality. Additional information from patent data can be used to overcome this hurdle. Industrial economists have documented a strong positive correlation between the value of a patent and elements such as the number of citations that the patent receives from subsequent research, its claims and scope.[2] Hence, the main indicator of patent quality used in this chapter is the number of citations that a patent receives, but alternative measures are used for robustness. The reduced form results indicate that the number of citations a patent receives decreases as firm size increases. This implies that firm size is negatively related to innovation quality. The use of the alternative proxies for innovation quality do not dramatically affect the results.

---

[2]Patent "scope" will be defined later in the chapter. For more details, see Lerner (1994).

The rest of the chapter is organized as follows: Section 2.2 reviews the related empirical literature. Section 2.3 describes the data. Section 2.4 presents the results of the analysis of firm size and growth. Section 2.5 focuses on the relationship between firm size and R&D intensity, while section 2.6 looks into the relation between patent quality and firm size. Finally, section 2.7 concludes.

2.2 Related Literature

The literature on firm growth has been deeply influenced by Gibrat's law. This theory has however been mostly refuted by empirical evidence, in particular for the US economy. Most studies showed that firm size in the U.S. is negatively correlated with firm growth (Mansfield, 1962; Hall, 1987; Amirkhalkhali and Mukhopadhyay, 1993; Botazzi and Secchi, 2003 for quoted US manufacturing firms and Evans, 1987a; Gabe and Kraybill, 2002 for non-quoted firms). Nevertheless, several studies have still argued that Gibrat's law holds for firms above a certain size threshold (Mowery, 1983; Caves, 1998).

Similarly, the early literature on firm level R&D has been inspired by Schumpeter's claim and showed that R&D spending increases more than proportionally in relation to firm size (Horowitz, 1962; Hamberg, 1964; Scherer, 1965 a,b, 1980; Comanor, 1967; Meisel and Lin, 1983; Kamien and Schwartz, 1982). However, these studies focused mainly on the 500-1000 largest firms, possibly due to limitations in data availability. In addition, the unobserved sector and firm heterogeneities, which were potential sources of bias, were not taken into account. Hence, it is not surprising that more recent studies from the late 80s and early 90s, which have utilized relatively larger datasets and controlled for sector level heterogeneity, have contradicted earlier findings. Indeed, they showed that the increase in firm size in relation to R&D is either proportional (Cohen et al, 1987) or slightly less than proportional below a certain firm size threshold (Bound et al, 1984).

The innovation patterns of different sized firms have been analysed. Kamien and Schwartz (1975) argue that small firms are more innovative in highly concentrated industries. Mansfield (1968) and Mansfield et al (1971) document that the major innovation per R&D expense is higher for smaller firms. Cooper (1964) claims that small firms have an organizational advantage in innovative activities. Other studies argue that major innovations come from small firms

because they make use of the innovation opportunities whereas large firms might supress such opportunities for various reasons (Blair, 1972; Pavitt and Wald, 1971; Kamien and Schwartz, 1975). These empirical studies are based on case studies which prevent us from drawing broader conclusions.

Finally the literature on patents has generated fruitful work for detecting firm level innovations and identifying their heterogenous values. Trajtenberg (1990) shows the positive relationship between the citations a patent receives and its social value. Harhoff et al (1996) shows the same relationship but for the private value of a patent. Hall et al (2001) argue that a firm's market value is not positively correlated with patent counts but rather with the "citation weighted patent portfolio", a measure which considers both patent counts and the number of citations per patent. Similarly, Shane and Klock (1997) document a positive relationship between patent citations and Tobin's q. In addition, Shane (1999a, 1999b) suggests that the more citations a patent receives, the more likely it is to get licensed. Sampat (1998) and Sampat and Ziedonis (2004) present the positive correlation between the number of citations and licensing revenues. This seems to confirm that the number of citations a patent receives is a good indicator of the quality of the innovation and justifies its use as a proxy for innovation quality.

In addition, other studies have introduced alternative indicators for innovation quality. Lerner (1994) shows that the scope of a patent is positively correlated with firm value. Lanjouw and Schankerman (1999, 2001) point out that the number of claims of a patent is correlated with its value. Jaffe and Trajtenberg (2002) argues that the number of claims of a patent could indicate the scope of a patent. For robustness, I verify my results using the number of claims of a patent as an alternative indicator of its quality. Finally, one could argue that a high-quality or more radical innovation would combine the knowledge from different technology fields. Therefore the originality index, defined by Trajtenberg, Jaffe and Henderson (1997) will be used as another alternative indicator for the quality of innovation. The details are provided later in the chapter.

2.3 Data

The data used is from the Standard and Poor's Compustat database which includes information on all publicly trade companies in the US since 1950. Besides the tremendous increase in the sample size, the average firm size in Compustat data has decreased drastically over time (see figure 1) which allows the sample data to capture a wider range of firm sizes. Compared to earlier studies using Compustat, this could potentially make the current sample more representative of the whole population of US manufacturing firms. Following a common practice in the literature, I exclude from the sample non-manufacturing firms or non-domestic firms.[3]

The data on innovation is taken from the NBER/USPTO patent dataset, described in detail in Hall et al (2001). It includes all utility patents granted in the USA between 1963 and 2002. By definition, patents grant their holder a monopoly for the use of the innovation. Patents identify "prior art" through *citations* so as to clearly mark the boundaries of that monopoly power. This variable, namely the number of citations will be used as a proxy of innovation quality. It is available in the NBER Patent Data set for all patents granted after 1975.

Another important component of the patent data set is the match between patent and Compustat firms (Hall et al, 2001). Even though this match is a great source for linking the patents to firms, it is necessary to be cautious because it includes the Compustat firms that existed in 1989. Firms that were established after 1989 are not included in the patent-match data.

The last two variables used are the *number of claims* and the *originality index*. The former is the number of "components" or the "main pieces" of the patent. It can be interpreted as the "scope" or "width" of the innovation (Jaffe and Trajtenberg, 2002, p. 432). The latter is an index calculated for each patent i as follows : $Originality_i \equiv 1 - \Sigma_j^{n_i} s_{ij}^2$, where s_{ij} is the percentage of citations made by patent i to a patent that is in patent class j and n_j is the number of patent classes. This implies that if a patent cites a wider set of patent classes, its originality index is higher.

Table 1 contains the descriptive statistics for the key variables.

---

[3] The term "non-domestic" refers to the firms that do not have their headquarters in the US.

2.4 Firm Size vs Firm Growth

2.4.1 Specification

To investigate the link between firm growth and firm size, the following regression is estimated using OLS

$$g_{it} = \beta_0 + \beta_1 \ln Sales_{it} + \delta_j + \delta_t + \varepsilon_{ijt} \tag{2.1}$$

where $g_{ijt} \equiv Sales_{ij(t+1)}/Sales_{ijt} - 1$ is the growth rate of firm i in industry j at time t, $Sales_{ijt}$ is net sales and δ_i and δ_t are respectively the industry and sector fixed effects.[4] This regression does not include the final period's growth rate of the firms that exit the sample. If Gibrat's law holds, β_1 should be zero.

The results are reported in the first column of Table 2. The coefficient β_1 is estimated as -0.037 with a standard deviation of 0.001 which is a considerable departure from Gibrat's law. This is both statistically significant and economically large. An increase of one standard deviation in $\ln Sales$ (= 0.412) decreases the growth rate by 1.5 percentage points. Since the average growth rate is 12 percent, this amounts to a sizeable 13 percent decrease in growth rate relative to the mean. However, simple OLS results may not be reliable enough and further refinements to the estimation are presented next.

Before the discussion of the results continues, it is important to note that the reduced form analysis on R&D intensity and innovation quality will use similar specifications to (2.1). To avoid repetition, the detailed discussion on the common caveats related to all three specifications will be provided in the first section and only be mentioned briefly in later sections.

2.4.2 Survival Bias

Firms that exit the market do not report any data for the year in which they exit the Compustat sample which disables the observation of the decline in $Sales_t$ in period t when the exit happens. In general, the exit rate is higher among smaller firms, so that excluding exiters from the sample biases the results in favor of small firms. To address this problem, there are two different methods.

---

[4] All nominal variables are deflated by the GDP deflator in the corresponding year

The first method consists in merely assuming that the output of a firm drops to zero right before the exit, so that the growth rate in their last period is -100 percent. The results of this method are presented in column 2 of Table 2. The parameter estimate has now increased from -0.037 to -0.025 confirming the suspicion that the benchmark OLS is biased downward due to survival bias. A one standard deviation increase in $\ln Sales$ leads to a 20 percent decrease relative to the average growth rate of this sample (0.057). Even though this method has generated both statistically and economically significant estimates, it is likely to overcorrect for the bias, since -100 percent is a lower bound for the actual growth rate in the last period before exit.

The second possible solution is to apply a Heckman two-step selection estimation (Wooldridge, 2000). The selection equation can be written as,

$$z_{it}^* = \gamma_0 + \gamma_1 \ln Sales_{it} + \boldsymbol{\gamma} \mathbf{X} + \epsilon_{it} \qquad \text{(Selection Equation)}$$

where z^* is a latent variable interpreted as "the propensity to exit", which depends on firm size and some other explanatory variables $\mathbf{X}$. Even though the latent variable is unobserved, a binary variable z_{it} is observable which is defined as:

$$z_{it} = \begin{array}{l} 1 \text{ if } z_{it}^* > 0, \text{(the firm remains in the sample)} \\ \\ 0 \text{ if } z_{it}^* < 0 \text{ (the firms exits)} \end{array}$$

The outcome equation is then

$$g_{it} = \beta_0 + \beta_1 \ln Sales_{it} + \delta_j + \delta_t + \varepsilon_{ijt} \qquad \text{(Outcome Equation)}$$

The growth rate g_{it} is observed only when $z_{it} = 1$. A consistent estimator of β_1 (2.1) can then be obtained from the conditional regression

$$\mathbb{E}\left(g_{it} \mid z_{it}^* > 0\right) = \beta_0 + \beta_1 \ln Sales_{it} + \beta_2 \lambda_{it} + \delta_j + \delta_t + \varepsilon_{ijt} \qquad (2.2)$$

where

$$\lambda_{it} = \frac{\phi\left(\gamma_0 + \gamma_1 \ln Sales_{it} + \boldsymbol{\gamma}\mathbf{X}\right)}{\Phi\left(\gamma_0 + \gamma_1 \ln Sales_{it} + \boldsymbol{\gamma}\mathbf{X}\right)}$$

is the "Inverse Mills Ratio". The parameters of $\gamma_0, \gamma_1, \boldsymbol{\gamma}$ are obtained with a probit regression. The Heckman two-step model requires an exclusion restriction (instrument) otherwise the

83

model is identified solely on distributional assumptions. The novel instrument used here is the ownership status of the firm. Firm exits are significantly affected by their ownership status. For example, subsidiaries would have a higher probability of dropping out of the Compustat sample. Indeed, if the parent company is financially constrained, it might decide to sell its subsidiary to a third party through carve-out, spin-off or sell-off (Draho, 2004, p.156). In addition, due to arbitrage, any additional growth opportunity in either the subsidiary or parent firm will be exploited by the parent company preventing any systematic relationship between firm status and growth. This fact makes the ownership status of the firm satisfy the exclusion restriction and, as result, a novel instrument in the selection equation.

Columns 3a,b and 4 of Table 2 report the estimation results of this two-step selection model. The significance of the inverse Mills ratio confirms the existence of a selection bias. Column 3a shows that the probability of surviving in the sample increases significantly in firm size. In addition, the sign of the coefficient on the ownership status dummy[5] implies that subsidiaries have a lower probability of survival. Column 3b reports the same results in Column 3a as the marginal effects at the sample means. These transformed results suggest that the marginal effect at the sample mean of $\ln Sales$ is 0.008 (standard error=0.000) which implies that the marginal change in $\ln Sales$ from the average increases the probability of surviving in the sample by 0.8 percent. The new, corrected coefficient on firm size in Column 4 is now -0.028 with a standard error of 0.002. As expected, it lies between the simple OLS coefficient and the coefficient obtained when assuming zero output in the last period. This is because, as explained, the OLS estimation favored small firms whereas the latter method favored the large firms. With this new corrected estimate, one standard deviation increase in $\ln Sales$ is associated with a 10 percent decrease in the growth rate, which is economically sizeable. Since the coefficient on firm size remains significantly negative, these results show that the survival bias alone cannot account for the negative relationship between firm size and growth.

2.4.3 Measurement Error

Another common concern for estimating (2.1) is the errors-in-variables problem, as firm sales are a noisy proxy for firm size. Consider a measurement error in $Sales_{it}$ in period t. This will

---

[5]This dummy becomes 1 when the observation is a subsidiary.

create a spurious negative relationship between firm growth and firm size because it affects the denominator of the dependent variable as well. Therefore the existence of a measurement error would bias the estimate of β_1 downwards. The problem can be solved through an instrumental variables approach, which uses the one year-lagged values of sales, $\ln Sales_{t-1}$ as an instrument for $\ln Sales_t$.

The results reported in column 5 of Table 2 show a less negative coefficient of -0.028 (standard deviation=0.001) on firm size, confirming the presence of some error-in-variables bias. Again, this suggests that one standard deviation increase in $\ln Sales_t$ is associated with 1.2 percentage points decrease in growth rate which corresponds to a 10 percent decrease of the growth rate relative to the sample average. Therefore this implies that errors-in-variables was not the sole driver of this relation.

2.4.4 Sample Selection

A final problem is the sample selection bias inherent in the Compustat sample. Suppose for example that most big firms are present in the sample, even though they might have low or medium productivities, but that only the small firms with the highest productivities are admitted into the sample. There are two potential reasons for this. First, small firms might be more productive due to some unique unobserved individual characteristics. Secondly, as part of their life-cycle, at the time that they are selected into the sample, they might be on the increasing side of an inverse U-shaped growth/age trajectory which might lead to high initial growth rate and lower growth rates over time.

These two issues require different approaches. The former concern is related to unobserved firm heterogeneities which can be captured by the fixed-effects regression. Column 6 of Table 2 reports the fixed effects results including time effects. Contrary to the expectations, adding firm fixed effects sharpens the contrast between the growth rates of small and large firms by reducing the coefficient down to -0.175 (standard error=0.003) which is more than 5-fold in absolute terms of the benchmark OLS estimate. This suggests that a one standard deviation increase in firm size is associated with an approximately 61 percent decrease in growth rate relative to the sample average. Even though this result seems to be surprising initially, it shows that avoiding the firm fixed effects generates an omitted variable bias which favors the large

85

firms. These omitted variables are likely to bias the results in favor of large firms, since these are the factors which have made these firms "large firms" at the first place. The very same factors would make the firm grow faster than others in general and not controlling for them would make the firm size take the credit for the additional growth of the large firms. Unfortunately, previous literature has generally overlooked this aspect, possibly due to data limitations.

The second issue regarding the position of the firm on its growth trajectory is related to the age or the time elapsed since when the firm first shows up in the sample. A way to address this issue is to control for the number of years that the firm has spent in Compustat sample. Column 7 controls for this variable which is denoted by "Age". The point estimate of the coeffcient on Age is -0.003 (standard error=0.000). This significant negative coefficient on Age illustrates that on average, firms are growing faster earlier in their life-cycle. This is also economically significant since this implies that in 10 years ($\cong$one standard deviation of Age, 11.3), the growth rate of a firm with the same beginning and end-of-period sizes will decrease by 25 percent relative to the average growth rate of the sample. On the other hand, the coefficient of the ln $Sales$ is close to the OLS estimate with a small increase. Controlling for age, the benchmark OLS estimate increases by only 0.007 percentage point to -0.03 (standard error=0.001), which is still highly significant and economically large. The tiny increase suggests that only a small portion of the OLS estimate could be attributed to the sample selection problem of the second kind.

Another strategy to address the second caveat is to focus only on the firms that have been long enough in the sample long enough. For that purpose Column 8 consider the firms that have been in the Compustat sample for more than 10 years. This approach is expected to eliminate the initial heterogeneity of the firms that might have qualified them to go public. The drawback of this approach is the significant decrease in the sample size (around 50%) and reduction of attention to healthy firms which manage to stay for more than 10 years in the sample. The average firm size increases by more than 15 folds (from 0.07 to 1.06) eliminating the young and typically small firms. Nevertheless, this method would be informative as regards the upper bound of the actual coefficient estimate on ln $Sales$. This specification produces an estimate of -0.011 (standard error=0.001) which is about a third of the OLS estimate. This decline was expected as the nature of the sample has changed, yet in spite of this, the result from this restricted sample still indicates a smaller but very significant negative relationship between

firm size and growth. Moreover, this negative relationship is also economically significant. A one standard deviation increase in ln $Sales$ ($\cong 2.154$) is associated with 2.37 percentage points decrease in the growth rate, which amounts to a 42 percent decrease in growth rate relative to the average growth rate of 5.6 in the sample.

Therefore the conclusion of all these different specifications is that the statistically and economically significant negative relationship between firm growth and firm size is robust to sample selection problem.

2.4.5 Additional Caveats

Time-varying Sector Characteristics. These heterogeneities could be a potential issue. The take-off of the computer industry in the mid-80s for example could be attributed to such an unobserved sector level change. In order to prevent the results from being affected by these changes, I introduce an interaction term between the sector and time dummies. Column 9 shows that the impact of firm size on growth is almost identical as in the benchmark OLS estimation, showing that time-varying sector effects do not significantly affect the benchmark results.

Short-Run vs Long-Run. The short-run relationship between firm size and growth might differ from the true long-run relationship , because of transitory and cyclical movements in Sales. Therefore, the 10-year average growth rate of firms between 1995-2005 is regressed on the initial firm size of 1995. This is done first by excluding firms which have exited over the period (Column 10) and then by applying Heckman's 2-step selection method (Columns 11a,b and 12). Column 10 indicates that the long-run results, which includes only the firms that survived for 10 years, are 0.009 percentage point higher than the short-run benchmark OLS. This should be expected since this sample excludes the exiters which are typically smaller firms. The average firm size of the survivors is 0.171 whereas the average firm size of the exiters during this period is -0.599. Since this might cause a survivor bias on the estimate, columns 11-12 employ the Heckman selection method. Column 11a shows that the probability of surviving is increasing in firm size whereas subsidiaries have a lower probability of surviving. Column 11b reports the marginal effects at the sample means. These estimates imply that the marginal deviation from the average firm size increases the probability of surviving by 5 percent. Similarly, being a subsidiary company decreases the probability of surviving by 41 percent.

However, the coefficient β_2 of the inverse Mills ratio in (2.2) is not significant indicating that the survival bias does not affect the OLS results systematically. Column 12 reports the results of the second step. The corrected estimate on $\ln Sales$ is now -0.026 (standard deviation=0.006). This is not very different from the OLS estimate in Column 10 since the survivor bias is not in effect. These alternative specifications indicate that even the long-run relationship between firm size and growth is not consistent with Gibrat's law.

Segment Level Analysis. Many Compustat firms are conglomerates of different business segments. Resource allocations might be different across the segments in a diversified firms. Therefore, the reader might wonder if the firm-level findings also hold at the segment level. To address this issue, I use Compustat segment files covering the period 1980-2005. For each segment, these files provide basic accounting information such as sales, assets, capital expenditures, operating profits, Standard Industrial Classification (SIC) codes starting in 1990 and onward. I rerun the benchmark regression of column 1 with the segment level data in column 13 where $\log(Sales)$ stands for the segment size. Compared to the benchmark case, the magnitude of the coefficient estimate becomes even bigger in absolute value (-0.046) showing that the significant negative relationship between firm size and firm growth hold at the segment level as well.

In conclusion, the benchmark regression, together with the robustness checks[6] show the following result.

Fact 1. *Firm size is negatively related to firm growth.*

2.5 Firm Size vs R&D Intensity

2.5.1 Specification.

The next empirical question is the relationship between firm size and R&D intensity. Bound et al (1984) had used a much older version of the Compustat data including only the firms that existed in 1976. With this cross-section, they found the size elasticity of R&D to be 0.97 suggesting that R&D intensity is independent of firm size.

---

[6] I also checked the robustness of the results with Compustat segment level data and I did not find any significant difference between the results of the two datasets.

To analyse the same relationship, I will consider the following benchmark specification,

$$\ln\left(R\&D_{it}/Sales_{it}\right) = \beta_0 + \beta_1 \ln\left(Sales_{it}\right) + \delta_j + \delta_t + \varepsilon_{ijt} \tag{2.3}$$

In this reduced form equation $R\&D_{it}$ stands for the total R&D spending of a company in a given year.

The results of this benchmark regression are in column 1 of Table 3 . In contrast to most previous studies, a 10 percent increase in firm size is associated with 2.65 percent decrease in R&D intensity. In other words, the elasticity of R&D spending with respect to the firm size is 0.735 which is significantly less than unity. To see the practical importance of this effect, consider an average sized firm ($\ln Sale = 0.073$) with the average R&D intensity ($\ln(R\&D/Sale)$ =-3.205). Consider also a 10 percent annual growth in real sales for 10 years. If the R&D intensity was independent of the firm size, the new real R&D spending would have increased from 4.4 to 11.4 percent at the end of the period. However, with the current estimate, the real R&D spending goes up to 8.9 percent which is 22 percent lower than the case where R&D intensity is independent of the firm size.

Unfortunately, about 30% of the observations are excluded due to missing entries, which is a problem that is addressed next.

2.5.2 Missing Observations

In the Compustat dataset, while most of the firms report detailed data, some firms report more aggregated variables.[7] As a result of these classification strategies, many firms do not report R&D, even though they might have had a positive R&D expenditure. Following Bound et al (1984), I will treat both missing and null observations as signifying "not reporting positive R&D" (see Bound et al (1984) for a detailed discussion of this assumption). Consequently, a two-step Heckman selection procedure is applied. The instrument that will be used in the first stage is the Propensity to Report (PTR) Index which is generated in the following way: First, 12 data items (excluding R&D) that are common to all the manufacturing firms are selected.[8]

---

[7] Firms have to follow certain accounting procedures in data reporting. However, the interpretation of these procedures vary in great detail across firms.

[8] The items are: DATA3 (inventories), DATA4 (current assets), DATA5 (current liabilities), DATA7 (property, plant & equipment), DATA29 (employees), DATA41 (cost of goods sold), DATA42 (labor related expenses), DATA103 (depreciation expenses), DATA216 (stockholders' equity), DATA224 (nonoperating income), DATA263

Reporting of these items substantially differs across firms. Hence, the PTR index of a firm in year t corresponds to the percentage of those 12 items that are reported (i.e. not missing) in that year. This index indicates to what extent and in how much detail a firm reports its data. Its value becomes 1 if the firm reports all of the 12 items and 0 if it reports none. The rationale behind this index as an instrument in the first stage is that, if the firm chooses to report a high fraction of those 12 items, then this would be a sign that the firm has a good propensity to report data, or detailed data classification. Hence, it is plausible to think that this firm will also be more likely to report R&D.

Table 3, columns 2a,b and 3 report the results of the selection model. The significance of the inverse Mills ratio indicates as before the existence of a selection bias. Column 2a shows that the probability of reporting positive R&D is increasing both in firm size and PTR index. Column 2b reports the marginal effects at the sample means. It shows that a marginal deviation from the sample average of firm size increases the probability of reporting positive R&D by 0.04 (standard error=0.001). The corrected estimates in Column 3 indicate that the R&D intensity decreases by 2.1 percent if the firm size increases by 10 percent, which corresponds to an elasticity of 79%, again significantly less than unity. To see its economic significance, the same exercise as above gives 19 percent lower real R&D spending in the final period relative to the case where the R&D intensity is independent of the firm size. Therefore the coefficient is still economically large.

2.5.3 Measurement Error

Similar to section 2.4.3, measurement errors could be a concern. I follow the same steps as before and instrument the sales at time t by its lagged value. The estimates, presented in Column 4 of Table 3, show the coefficient on ln *Sales* is precisely estimated as -0.225 (standard error=0.003). In other words, the results were not strongly biased by the measurement error, since the R&D elasticity increases from 74 percent to only 78 percent.

(building at cost), DATA264 (machinery at cost).

2.5.4 Sample Selection

This problem has been discussed in the previous section in detail. Briefly, if the Compustat sample selects only the highest R&D intensive firms among the small firms, and if this bias is not so severe for large firms, this factor might derive the observed results in the benchmark OLS. Therefore I will address this issue again in two ways.

First, the permanent unobserved heterogeneity is captured through fixed effects. The results of this method are presented in column 5. Interestingly, the coefficient estimate of this specification is lower than the benchmark OLS estimate, namely -0.384 with a standard error of 0.007, and the size elasticity of R&D is now only 62% , marking even more the contrast to the previous literature. The possible explanation for this result is that the unobserved firm characteristics were affecting R&D intensity and size positively, thus biasing the coefficient on firm size upward , if omitted. For example firm specific know-how or products could affect the size of the firm through making it capture the market and grow faster. The very same factors could make the firm invest in R&D to develop them further.

Secondly, the temporary unobserved heterogeneity caused by the firm's life-cycle can be captured by the number of years spent in the Compustat sample. The results from this estimation, in Column 6 confirm that firms are more R&D intensive when recently added to the sample, but this does not have a dramatic effect on the benchmark estimates. To address the same problem in a different way, one can restrict the sample to only those firms which have been surveyed by Compustat for more than 10 years. The resulting estimate in Column 7 is now -0.092, which is higher than the benchmark. As it has been mentioned in the previous section, the possible explanation for this result is the exclusion of small firms from the sample. However, even among mature firms, the negative relationship between R&D intensity and firm size remains, though less stark. Consider a firm that has the average size and R&D intensity in this new sample (ln $Sale$ =1.027, ln($R\&D/Sale$ =-3.705) and grows with 10 percent for 10 years. The end-of-period real R&D spending was going to be 0.028 if R&D intensity was independent of the firm size, but is now only 0.021, that is, 8% lower.

2.5.5 Additional Caveats

The robustness of the results to alternative specifications can also be checked. The methods are the same as the ones employed in section 2.4.5.

Changing Sector Heterogeneities. The results using a year×sector interaction effect are presented in column 8 and are very close to the benchmark OLS results.

Short-Run vs Long-Run. The long-run estimate of -0.298 (standard error=0.014) in column 9 is 10% lower than the short-run estimate. This coefficient denotes that the elasticity of R&D with respect to the firm sale is 0.70. As a result, the negative relationship between R&D intensity and firm size is sharper in the long-run.

Segment Level Analysis. Column 10 reruns the benchmark regression using segment level data. The new estimate of -0.353 shows that the negative relationship between firm size and R&D intensity is stronger at the segment level.

Overall, the following fact can be summarized from the data.[9]

Fact 2. *Firm size is negatively related to R&D intensity.*

2.6 Firm Size vs Innovation Quality

2.6.1 Specification

The final and most innovative reduced form estimated is the one between firm size and innovation quality, as proxied by the number of forward citations. The benchmark reduced form regression considered is:

$$Citations_{jit} = \beta_0 + \beta_1 \ln(Sales_{jt}) + \delta_k + \delta_t + \varepsilon_{ji} \tag{2.4}$$

where $Citations_{jit}$ denotes the number of citations that patent i of firm j receives within 7-years after its grant date t, $Sales_{jt}$ is the sales of firm j in year t, and δ_k is a fixed effect for 4-digit International Patent Classification (IPC) k. In chosing the window size, one faces a trade-off because a larger window captures more citations per patent, but also reduces the number of observations available. Since the number of cumulative citations increases mostly until 5 to 6

---

[9]I also checked the robustness of the results with Compustat segment level data and I did not find any significant difference between the results of the two datasets.

years before evening out, after the grant date, using a 7-year period seems optimal.[10]

The first column in Table 4A reports the benchmark result, using the full sample of patents and firms. The coefficient estimate is precisely estimated as 0.154 (standard error = 0.008). However, there might be a sample bias. The patent data sample was matched only to firms which existed in 1989. Hence, the firms that entered after 1989 are not represented. This causes the sample to get smaller, older and the average firm size to get larger over time; hence, the sample could get biased as time elapses.

To account for this, the second column focuses only on patents obtained between 1988-1992. In this case, the benchmark coefficient doubles to -0.324. This result should not be surprising because the previous sample included presumably large and successful firms from later years in addition to the current sample, which would bias the result in favor of large firms. This estimate is also economically sizable. A one-standard deviation increase in the log of firm sale (= 2.294) is associated with a decrease of 0.742 citations, which is 15 percent of the sample mean.

In columns (3a) and (3b) a negative binomial regression is performed. The reason for using this form rather than a Poisson model is to allow for the overdispersion in citations, i.e., that the standard deviation of citations is greater than the mean. The coefficients of the negative binomial and the marginal effects at the sample averages are reported respectively in columns 3a and 3b of the same table. The negative binomial estimate at the sample mean (= -0.282) is slightly larger than the OLS estimate (= -0.324) and still highly significant (standard error = 0.013). In column 4 the analysis is extended to a zero-inflated negative binomial to correct for the fact that a lot of observations cluster around zero.[11] The coefficient estimate at the sample average (= -0.276) does not differ much from the standard negative binomial. Hence, it appears that the patents of small firms receive more citations and hence that smaller firms generate better quality patents.

However, there remain some important issues to be addressed.

---

[10] I checked for the robustness of the results with 8 and 10-year windows and verified that the results are not significantly sensitive to those variations.

[11] a) In the full sample 40%, in the 88-92 sample 12% of the patents have 0 citations.

b) The independent variables of the inflation equation are the firm sales and 2-digit IPC classifications.

2.6.2 Patenting Decisions

An important concern is the differential patenting behavior of small and large firms. One could argue that due to strategic reasons large firms might patent more heavily. Indeed, one can consider that there is a "crown" patent, which is the main patent for an important innovation, but that large firms also patent other peripheral innovations around it, so that those can be used against opponents in cases of litigation, to make a strong case against a competitor coming even close to the market niche of the firm. On the other hand, if the costs of patenting are important to small firms, they might register their innovations only if the innovation is significantly high quality and prefer to save on the expense of patenting minor quality patents. This would cause the quality of patents of small firms to be higher. This concern can be addressed by performing the regression not on the whole universe of patents, but rather on only the best patent (the one with the highest number of citations) of each company. However, this strategy tends to overestimate the quality of patents of larger firms that apply for a big number of patents.[12] For that purpose, I will use the total number of patents of a firm as a control variable.

Column 5 shows the results of this approach. The results confirm the existence of a possible bias since the coefficient of the negative binomial regression has increased from -0.040 to -0.035. The marginal effect of this specification at the sample average is almost identical to the previous specifications. As a result, it is reassuring to see that the bias due to patenting behavior of firms is not the sole driver of the negative relationship between firm size and innovation quality.

2.6.3 Non-patenting Firms

The analysis so far have focused only on firms which hold at least one patent in their portfolio. However, 25% of the Compustat firms do not register for any patent in any given year. These firms are typically small firms, so that excluding them from the analysis could bias the results in favor of small firms. To address this issue, a Heckman selection model is utilized. The variable used here as an instrument is a dummy, taking the value of one if the firm has patented at all in the past. The rationale behind this is that a patent made before signifies on the one hand that the firm is already familiar with the patenting procedures, which makes it easier to patent

---

[12]Assume firm i receives only 1 patent in a given patent class in year t and firm j obtains 10 patents in the same patent class and year. If we focus on the best patents, firm j will be more advantegous since its best patent will be the best of 10 patents whereas firm i has only one patent which will count as the best patent.

new innovations and on the other hand, reveals that the firm has a propensity to patent its innovations (as opposed to some firms, which for secrecy or other reasons might chose to never patent their innovations). Therefore one could expect a higher probability of observing a firm in the sample if the firm has ever patented before. The results of this technique are listed in columns 6-7. The estimates of the first stage show that the probability of registering for a patent is increasing in firm size. In line with the expectations, the probability of registering for a patent increases if the firm has patented before. The significant inverse Mills ratio documents the existence of selection in the estimation. The corrected results show that the negative relation between firm size and patent quality remains, though the coefficient estimate ($= -0.242$) is bigger than the benchmark estimate ($= -0.324$) since the latter was biased due to the missing observations. Hence, the significant negative relationship between firm size and innovation quality is not solely driven by sample selection.

2.6.4 Segment Level Analysis

In this section, I match the patents to the segments of the firm. First I use the NBER's routine to match the patents to the Compustat firms. Next, I use the concordance of Silverman (2002) to identify the relevant SIC codes for each individual patent. Finally, within each firm, I distribute the patents to the relevant segments according to the the SIC code matches between patents and segments. Columns 8-10b repeat the exercises in columns 1-3b using segment data. The results show a stronger negative relationship between segment size and innovation quality. In the benchmark OLS case, the coefficient went down from -0.154 to -0.233. Similarly, the marginal effect went down from -0.282 to -0.601 when we restrict the sample to 1988-1992.

2.6.5 Alternative Quality Measures

Next, I consider alternative indicators of patent quality. Column 11 considers an originality index, the construction of which was explained in section 2.3 ranging from 0 to 1, and increasing in the width of technologies cited by the patent. For any given number of citations, the wider is the range of cited patent classes, the higher the value of this index will be.[13] The results in

---

[13] The calculation of this index was described in the data description section.

95

column 8 indicates that as the firm size doubles, its innovations combine a smaller number of technologies and the originality index decreases by 0.001.

In column 12 the results from regressing the number of claims of a patent on the firm size are presented. The number of claims could be seen as an alternative indicator of the value of an innovation since the patent holder would demand so many claims that the marginal cost is equated to the marginal expected return to that innovation. The result of this estimation also documents a negative relationship between firm size and innovation quality. The coefficient estimate of this regression ($= -0.601$) is not only statistically but also economically significant. A one-standard deviation increase in firm size ($= 2.294$) is associated with a decrease of 1.378 claims, which is 11 percent of the sample mean.

In conclusion, I proxied the innovation quality by patent citations, by patent originality and by the number of patent claims and all these cases generated the following result.[14]

Fact 3. *Firm size is negatively related to innovation quality.*

2.7 Conclusion

This chapter investigated the relationship between firm size and three firm characteristics: firm growth, R&D intensity and innovation quality. The first two of these characteristics have received a lot of attention from the literature since the early half of the 20th century. The empirical findings had been inconclusive on the subject. Therefore, this chapter has generated its own independent results using a recently updated version of the Compustat data. In addition, it has addressed several important caveats previously overlooked. The two of the main results are that firm size is negatively related to R&D intensity and sales growth.

The third major and most innovative question in this chapter is the relationship between firm size and innovation quality. Previous studies have been constrained due to the lack of a proxy for innovation quality. This chapter made use of the strong empirical evidence on the close relationship between patent citations and innovation quality and proxied quality with citations. For robustness checks, alternative proxies have also been utilized, such as number of patent claims and an originality index. The key finding of all these reduced form analysis

---

[14]Controlling for the number of years spent in the Compustat sample and restricting attention only to 10+ years old firms did not change the results dramatically. Therefore they are not being reported to save space.

was that smaller firms produce higher quality innovations. This result was robust to sample selection and firms' patenting decisions.

The empirical results in this chapter should form the basis for a theoretical analysis on the subject. It represents a first important step towards recognizing the heterogeneous contributions of different-quality innovations. However, there many important questions remaining: Does every firm make the same type of innovation? Do small firms mainly invest in product innovation whereas large firms in process innovation? Do firms invest more in cost-cutting process innovations during recessions and product innovations during economic booms? These are only a couple of very intriguing questions with important policy implications. The next step of this research agenda will be to identify innovations as product versus process innovations and to address these open questions.

2.8 Appendix A: Tables & Figures

Figure 1
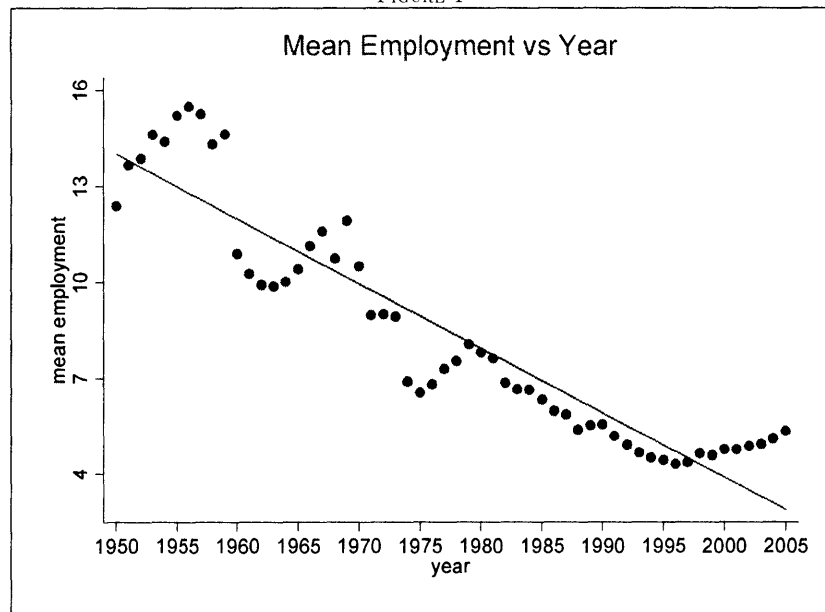
TABLE 1. SUMMARY STATISTICS

Variable	Obs	Mean	St Dev
Growth_t	111,869	0.118	0.412
$\text{Log(Sales}_t)^*$	118,343	0.073	2.294
$\text{Log(R\&D}_t)^*$	64,214	-3.228	2.231
Age	118,343	12.012	11.329
Subsidiary Dummy	118,343	0.019	0.134
PTR Index	118,343	0.744	0.140
Citations - 7 year	346,719	5.077	7.402
Originality	346,719	0.348	0.289
Claims	346,719	12.743	10.914

*Sale and R&D data is normalized by the GDP deflator

99

Variable	OLS (1)	OLS (-100%) (2)	Heckman 2-Step survive (3a)	mfx (3b)	growth (4)	IV (5)	OLS (6)
Log(Sales$_t$)	-0.037***	-0.025***	0.076***	0.008***	-0.028***	-0.028***	-0.175***
	(0.001)	(0.001)	(0.003)	(0.000)	(0.002)	(0.001)	(0.003)
Subsidiary			-0.422***	-0.062***			
Dummy			(0.046)	(0.009)			
Constant	0.331***	0.235***	0.6.583***		0.304***	0.241***	0.008***
	(0.026)	(0.030)	(0.106)		(0.052)	(0.022)	(0.014)
$\lambda-$ Inv Mills Ratio					0.666***		
					(0.098)		
Sector Dummy	Yes	Yes	Yes		Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes		Yes	Yes	Yes
Firm Dummy	No	No	No		No	No	Yes
R^2	0.067	0.040				0.051	0.297
Obs	111,869	118,343	118,343 (6588 censored)			102,532	111,755

Note. The technique used for each estimation is reported on top of its column

Dependent variable is sales growth=Sale(t+1)/Sale(t) -1.

"mfx" indicates the marginal effects at the sample mean.

Heteroskedasticity robust standard errors are reported in paranthesis

*, **, *** indicate 10%, 5%, 1% significance

TABLE 2B. GROWTH REGRESSIONS

Variable	OLS (7)	OLS (Age>10) (8)	OLS (9)	OLS LR (10)
Log(Sales$_t$)	-0.030***	-0.011***	-0.037***	-0.028***
	(0.001)	(0.001)	(0.001)	(0.002)
Age	-0.003***			
	(0.000)			
Subsidiary Dummy				
Constant	0.222***	0.154***	0.124***	0.217***
	(0.015)	(0.013)	(0.001)	(0.024)
Sector D.	Yes	Yes	No	Yes
Year D.	Yes	Yes	No	No
Firm D.	No	No	No	No
Year*Sector D.	No	No	Yes	No
R^2	0.072	0.044	0.044	0.221
Obs	111,755	47,942	111,755	1,489

Note The technique used for each estimation is reported on top of its column

Dependent variable is sales growth=Sale(t+1)/Sale(t) -1

"mfx" indicates the marginal effects at the sample mean

Heteroskedasticity robust standard errors are reported in paranthesis.

*, **, *** indicate 10%, 5%, 1% significance

TABLE 2C. GROWTH REGRESSIONS

Variable	Heckman 2-Step			OLS
	LR, survive	LR, mfx	LR, growth	(segment)
	(11a)	(11b)	(12)	(13)
Log(Sales$_t$)	0.116***	0.046***	-0.026***	-0.046***
	(0.010)	(0.004)	(0.006)	(0.001)
Age				
Subsidiary	-1.605***	-0.412***		
Dummy	(0.373)	(0.035)		
Constant	0.492**		0.191***	0.243***
	(0.234)		(0.142)	(0.027)
Sector D.	Yes		Yes	Yes
Year D.	No		No	Yes
Firm D.	No		No	No
Year*Sector D.	No		No	No
R^2				0.066
Obs	3,289 (1800 censored)			117,853

Note: The technique used for each estimation is reported on top of its column.

Dependent variable is sales growth=Sale(t+1)/Sale(t) -1

"mfx" indicates the marginal effects at the sample mean

Heteroskedasticity robust standard errors are reported in paranthesis.

*, **, *** indicate 10%, 5%, 1% significance.

TABLE 3A. R&D INTENSITY REGRESSIONS

Variable	OLS (1)	Heckman 2-Step report (2a)	mfx (2b)	log(R&D/Sales$_t$) (3)	IV (4)
Log(Sales$_t$)	-0.265***	0.095***	0.038***	-0.214***	-0.225***
	(0.003)	(0.002)	(0.001)	(0.004)	(0.003)
PTR Index		1.460***	0.581***		
		(0.048)	(0.019)		
Constant	-4.307***	-3.115***		-7.146***	-4.491***
	(0.157)	(0.099)		(0.275)	(0.157)
$\lambda-$ Inv Mills Ratio				1.126***	
				(0.069)	
Sector Dummy	Yes	Yes		Yes	Yes
Year Dummy	Yes	Yes		Yes	Yes
R^2	0.465				0.443
Obs	65,230	120,238 (55,008 cnsrd)			60,710

Note: The technique used for each estimation is reported on top of its column

Dependent variable is log R&D intensity= ln(R&D(t)/Sale(t)).

"mfx" indicates the marginal effects at the sample mean.

Heteroskedasticity robust standard errors are reported in paranthesis

*, **, *** indicate 10%, 5%, 1% significance

TABLE 3B. R&D REGRESSIONS

Variable	OLS (5)	OLS (6)	OLS (Age>10) (7)	OLS (8)	OLS LR (9)	OLS (segment) (10)
Log(Sales$_t$)	-0.384***	-0.248***	-0.092***	-0.257***	-0.298***	-0.353***
	(0.007)	(0.003)	(0.004)	(0.003)	(0.014)	(0.003)
Age		-0.006***				
		(0.001)				
Constant	-4.560***	-4.324***	-3.909***	-3.227***	-1.594***	-1.976***
	(0.148)	(0.155)	(0.098)	(0.005)	(0.280)	(0.099)
Sector Dummy	Yes	Yes	Yes	No	Yes	Yes
Year Dummy	Yes	Yes	Yes	No	No	Yes
Firm Dummy	Yes	No	No	No	No	No
Year*Sector D.	No	No	No	Yes	No	No
R^2	0.881	0.466	0.390	0.483	0.456	0.488
Obs	65,230	65,230	30,831	65,230	2,418	44,014

Note. The technique used for each estimation is reported on top of its column

Dependent variable is log R&D intensity= ln(R&D(t)/Sale(t))

"mfx" indicates the marginal effects at the sample mean.

Heteroskedasticity robust standard errors are reported in paranthesis

*, **, *** indicate 10%, 5%, 1% significance

Variable	OLS (75-95) (1)	OLS (88-92) (2)	Negative Binomial (88-92) (3a)	mfx, (88-92) (3b)	0-Inf'd Neg Binomial (88-92) (4a)	mfx, (88-92) (4b)
Log(Sales$_t$)	-0.154***	-0.324***	-0.045***	-0.282***	-0.042***	-0.276***
	(0.008)	(0.018)	(0.002)	(0.013)	(0.002)	(0.013)
Constant	3.119***	8.146***	2.425***		1.623***	
	(0.137)	(0.080)	(0.012)		(0.035)	
IPC4 Dummy	Yes	Yes	Yes		Yes	
Year Dummy	Yes	Yes	Yes		Yes	
R^2	0.116	0.142				
Obs	346,719	91,366	91,366		91,366 (11,208-zero obs)	

Note The technique used for each estimation is reported on top of its column

Dependent variable is the number of citation received by each patent

"mfx' indicates the marginal effects at the sample mean

Heteroskedasticity robust standard errors are reported in paranthesis.

*, **, *** indicate 10%, 5%, 1% significance

TABLE 4B. PATENT CITATION REGRESSIONS

Variable	Neg Binomial		Heckman 2-Step		
	Best, (88-92)	mfx	Selection	mfx	All firms, (88-92)
	(5a)	(5b)	(6a)	(6b)	(7)
Log(Sales$_t$)	-0.035***	-0.280***	0.315***	0.021***	-0.242***
	(0.003)	(0.023)	(0.003)	(0.000)	(0.028)
Count	0.067***	0.567***			
	(0.001)	(0.012)			
Pat Dummy			0.303***	0.016***	
			(0.030)	(0.001)	
$\lambda-$ Inv Mills Ratio			1.576***		
			(0.347)		
Constant	-17.890***		0.647***		1.268***
	(0.083)		(0.010)		(0.023)
IPC4 Dummy	Yes		Yes		Yes
Year Dummy	Yes		Yes		Yes
R^2					
Obs	31,176		97,999 (6,633 cnsrd)		

Note The technique used for each estimation is reported on top of its column

Dependent variable is the number of citation received by each patent

"mfx" indicates the marginal effects at the sample mean

Heteroskedasticity robust standard errors are reported in paranthesis

*, **, *** indicate 10%. 5%, 1% significance

TABLE 4C. PATENT CITATION REGRESSIONS

Variable	OLS segment, (80-95) (8)	OLS segment, (88-92) (9)	Negative Binomial		OLS Originality (11)	OLS Claims (12)
			segment, (10a)	mfx, (88-92) (10b)		
			(88-92)			
$\text{Log(Sales}_t)$	-0.233***	-0.440***	-0.057***	-0.601***	-0.001***	-0.601***
	(0.018)	(0.034)	(0.004)	(0.021)	(0.000)	(0.021)
Constant	-5.497***	8.344***	-2.165***		-0.299***	16.297***
	(0.738)	(0.969)	(0.111)		(0.002)	(0.120)
IPC4 Dummy	Yes	Yes	Yes		Yes	Yes
Year Dummy	Yes	Yes	Yes		Yes	Yes
R^2	0.057	0.047			0.111	0.050
Obs	132,029	52,623	52,623		30,865	92,718

Note The technique used for each estimation is reported on top of its column

Dependent variable is the number of citation received by each patent.

"mfx" indicates the marginal effects at the sample mean

Heteroskedasticity robust standard errors are reported in paranthesis.

*, **, *** indicate 10%, 5%, 1% significance

2.9 References

Amirkhalkhali, Saleh and Arun K. Mukhopadhyay (1993) "The Influence of Size and R&D on the Growth of Firms in the U.S" *Eastern Economic Journal*, 19(2), 223-233.

Bottazzi, Giulio and Angelo Secchi (2003) "Sectoral Specifities in the Dynamics of U.S. Manufacturing Firms" *LEM Papers Series*, 2003/18.

Bound, John, Clint Cummins, Zvi Griliches, Bronwyn H. Hall and Adam B. Jaffe (1984) "Who Does R&D and Who Patents?" in *R&D, Patents and Productivity*. Ed. Z. Griliches, University of Chicago Press, Chicago.

Caves, Richard E. (1998) "Industrial Organization and New Findings on the Turnover and Mobility of Firms" *Journal of Economie Literature*, 36, 1947-1982.

Cohen, Wesley M.,Richard C. Levin and David C. Mowery (1987) "Firm Size and R&D Intensity: A Re-Examination" *Journal of Industrial Economics*, 4, 543-565.

Comanor, William S.(1967) "Market Structure, Product Differentiation and Industrial Research" *Quarterly Journal of Economics*, 81, 639-51.

Cooper, Arnold.C. (1964) "R&D is More Efficient in Small Companies" *Harvard Business Review*, 42(3), 75-83.

Draho, Jason (2004) *The IPO Decision: Why And How Companies Go Public*, Edward Elgar Publishing, Northampton, MA

Evans, David S.(1987a) "The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries" *Journal of Industrial Economics*, 35(4), 567-581.

Gabe, Todd M. and David S. Kraybill (2002) "The Effect of State Economic Development Incentives on Employment Growth of Establishments" *Journal of Regional Science*, 42, 703-730.

Gibrat, Robert (1931) *Les Inégalités Économiques*, Librairie du Recueil Sirey, Paris.

Hall, Bronwyn H.(1987) "The Relationship between Firm Size and Firm Growth in the U.S. Manufacturing Sector," *Journal of Industrial Economics*, 35, 583-606.

Hall, Bronwyn H., Adam B. Jaffe and Manuel Trajtenberg (2001) "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools" *NBER Working Paper*, No. 8498. Reprinted as Chapter 13 in Jaffe and Trajtenberg (2002).

Hamberg, Daniel (1964) "Size of Firm, Oligopoly and Research: The Evidence" *Canadian Journal of Economics and Political Science*, 30, 62-7

Harhoff, Dietmar, Francis Narin, F. M. Scherer and Katrin Vopel (1999) "Citation Frequency and the Value of Patented Inventions" *The Reviewof Economics and Statistics*, 81(3), 511-515.

Horowitz, Irving L. (1962) " Firm Size and Research Activity" *Southern Economic Journal*, 28, 298-301.

Jaffe, Adam B., Manuel Trajtenberg and Rebecca Henderson (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations" *The Quarterly Journal of Economics*, 108(3), 577-598.

Jaffe, Adam B. and Manuel Trajtenberg (2002) *Patents, Citations and Innovations: A Window on the Knowledge Economy*, Cambridge: MIT Press.

Jaffe,Adam B. (2003) "Real Effects of Academic Research." *The American Economic Review* 79, 957-70.

Jones, Charles I. and John C. Williams (1998) "Measuring the Social Return to R&D" *Quarterly Journal of Economics*, 113(4), 1119-1135.

Judd, Kenneth L. (1998) *Numerical Methods in Economics*, MIT Press, Cambridge MA.

Kamien,Morton I. and Nancy Lou Schwartz (1975) "Market Structure and Innovation: A Survey" *Journal of Economic Literature*, 13(1), 1-37.

Lanjouw, Jean O. and Mark A. Schankerman (1997), "Stylized Facts of Patent Litigation: Value, Scope and Ownership" *NBER Working Paper*, No. 6297.

Lanjouw, Jean O. and Mark A. Schankerman (1999) "The Quality of Ideas: Measuring Innovation with Multiple Indicators" *NBER Working Paper*, No. W7345.

Lanjouw, Jean O. and Mark A. Schankerman (2001) "Characteristics of Patent Litigation: A Window on Competition" *RAND Journal of Economics*, 129-151.

Lerner, Joshua (1994) "The Importance of Patent Scope: An Empirical Analysis" *RAND Journal of Economics*, 25 (2), 319-333.

Mansfield, Edwin (1962) " Entry, Gibrat's Law, Innovation and the Growth of Firms" *American Economic Review*, 52(5), 1023-1051.

Mansfield, Edwin (1968) *Industrial Research and Technological Innovation: an Eco-*

nomic Analysis, Norton, New York.

Mansfield, Edwin, R. John, S. Jerome, W. Samuel and H. Michael (1971) *Research and Innovation in the Modern Corporation*, W.W. Norton & Company, New York.

Meisel, John B. and Steven A. Y. Lin (1983) "The Impact of Market. Structure on the Firm's Allocation of Resources to. Research and Development" *Quarterly Review of Economics and Business*, 23, 28-43.

Mowery, David C. (1983) "Industrial Research and Firm Size, Survival, and Growth in American Manufacturing, 1921-1946: An assessment" *Journal of Economic History*, 43, 953-980.

Pavit, Keith and S Wald (1971) *Conditions for Success in Technological Innovations*, OECD, Paris.

Sampat, Bhaven N.(1998) *Cite-Seeing: Patent Citations and Economic Value*, Colombia University, New York.

Sampat, Bhaven N. and Arvids A. Ziedonis (2004) "Patent Citations and the Economic Value of Patents" in Schmoch, Ulrich et al eds., *The Handbook of Quantitative Science and Technology*, 277-298, Springer.

Scherer, Frederic M.(1965a) "Firm Size, Market Structure, Opportunity and the Output of Patented Inventions" *American Economic Review*, 55, 1097-1125.

Scherer, Frederic M.(1965b) "Size of Firm, Oligopoly and Research: A Comment" *Canadian Journal of Economics and Political Science*, 31, 256-266.

Scherer, Frederic M.(1980) *Industrial Market Structure and Economic Performance*, Rand McNally, Canada.

Schumpeter, Joseph A. (1934) *The Theory of Economic Development: An inquiry into profits, capital, credit, interest and the business cycle*, Transaction Publishers, New Jersey.

Schumpeter, Joseph A. (1942) *Capitalism, Socialism and Democracy*, Routledge, London.

Shane, Hilary and Mark Klock (1997) "The Relation Between Patent Citations and Tobin's q in the Semiconductor Industry" *Review of Quantitative Finance and Accounting*, 9(2).

Shane, Scott (1999a) "Selling University Technology" mimeo.

Shane, Scott (1999b) "Technological Opportunities and New Firm Creation" mimeo.

Trajtenberg, Manuel (1990) "A Penny for Your Quotes: Patent Citations and the Value of Innovations," *RAND Journal of Economics*, 21, 172-187.

Wooldridge, Jeffrey M. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge MA.

Chapter 3

Firm Size, Innovation Dynamics and Growth: Theory and Policy Analysis

3.1 Introduction

Chapter 2 documented three important stylized facts in the data: Small firms grow faster, small firms are more R&D intensive and small firms produce higher quality innovations. On the theoretical side, much of the previous literature on innovation has focused mainly on innovation frequency with an economy-wide uniform innovation quality. Clearly an approach that overlooks the heterogeneity in the innovation quality of different-sized firms would not match the empirical facts and underestimate the potential impact of the radically innovating firms. As a result, this approach could lead to wrong policy conclusions.

The main questions of this chapter are: What are the theoretical micro-foundations behind the observed empirical patterns in the data? How is firm size related to the innovation quality? What is the implication of this relationship for firm growth? How should the optimal R&D policy incorporate these facts? These questions are at the heart of any policy debate which targets technological development through innovation.

To address these questions, first, a theoretical model is outlined. Next, the structural parameters of the theoretical model are estimated using the Simulated Methods of Moments (SMM). Finally, a policy experiment is conducted, in which different R&D subsidies are applied to different sized firms.

More precisely, in order to understand the microfoundations underlying the stylized facts, I build a general equilibrium model with entry and exit. In this model, technologically heteroge-

neous firms compete for innovation in order to increase operating profits. Each firm is identified by its firm-specific labor productivity which is improved stochastically through R&D and innovation. Each innovation raises the firm's profit, by reducing its cost relative to its competitors. A novelty of this model is that firms can endogenously choose not only the probability of innovation, but also the innovation quality which had been exogenous in previous papers. I prove the existence and characterize the properties of the equilibrium and show that the model's predictions are consistent with the aforementioned reduced form evidences. My theoretical results rely on two features: 1) The concavity of the profit function and, consequently of the value function; 2) R&D production technology being constant returns to scale in productivity. Any model with these features would imply that firms with lower productivity will have a greater incentive to increase their productivity. Therefore, they will be more R&D intensive, choose higher quality innovations and grow faster.

Next, I estimate the main structural parameters of the theoretical model using the Simulated Method of Moments (SMM). These parameter estimates allow me to discuss the impact of various macroeconomic policies on innovation. To complete the analysis, I focus on the effects of a public R&D subsidy for different sized firms by comparing three different regimes: i) no-subsidy, ii) uniform (size-independent) R&D subsidy, iii) size-dependent R&D subsidy. The results of this analysis document the significant contributions of subsidies on welfare. More interestingly, the optimal size-dependent policy provides higher subsidies to smaller firms since the marginal product of technology in final good sector is higher for small firms.

This chapter is a contribution to the endogenous growth literature with step-by-step innovation. Previous studies with endogenous R&D decisions have mainly focused on the arrival rate of innovations as the choice variable of firms, treating innovation size as exogenous (Aghion, Harris and Vickers, 1997; Segerstrom 1998; Aghion Harris, Howitt and Vickers, 2001; Acemoglu and Akcigit, 2008). This theoretical specification leaves out a very important empirical observation; namely that different innovations have different quality and technological impacts. The ultimate goal of the current work is to shed light on this unexplored subject.

The results of this chapter are also interesting for strand of literature focusing on R&D subsidies. It is widely argued that the social return to R&D is greater than the private return, which suggests that the equilibrium amount of private R&D investment is suboptimal and

R&D subsidies should be used to compensate for the private underinvestment (Spence, 1984; Griliches, 1992, 1995; Jones and Williams, 1998). In the theoretical model of this chapter, there will be several sources of positive externalities arising from technological innovation. Since these effects are not internalized by the firms, the equilibrium R&D investment will be below the socially optimal level. This chapter contributes to the R&D literature by introducing a size-dependent R&D subsidy policy and documenting the substantial welfare gain from this policy.

Finally this work is also a contribution to the recently growing literature on size-dependent policies. This literature typically argues that the size-dependent policies create detrimental effects on the economy by misallocating the resources (Guner, Ventura and Xu, 2008; Klenow and Rodriguez-Clare 1997; Hall and Jones 1999; Caselli 2004, Restuccia and Rogerson, forthcoming). On the other hand, in the context of Intellectual Property Rights Protection, Acemoglu and Akcigit (2008) argue that the gain from size-dependent policies could be substantial.

The rest of the chapter is organized as follows: Section 3.2 presents a simplified version of the main model; section 3.3 builds the main theoretical model and presents its results; section 3.4 structurally estimates the model with SMM; section 3.5 investigates the optimal size-dependent R&D subsidy policy and section 3.6 concludes.

3.2 Toy Model

This section presents a simplified static version of the main model in order to give some insights about the mechanisms at play.

3.2.1 Environment and Equilibrium

Consider the following static model with a final good sector and its production technology of the following form

$$Y = \left[A \sum_{i=1}^{N} y_i - \frac{\gamma}{2} \sum_{i=1}^{N} y_i^2 \right]^{1/2} \tag{3.1}$$

where y_i denotes the quantity of intermediate good i. $A > 0$ stands for the total factor productivity in the final good sector and is taken as exogenous. For simplicity, I assume that the number of the firms is constant and equal to $N >> 0$. In this simplified version, I abstract from

entry and exit. Each intermediate good is produced by a monopolist i with a linear production technology

$$y_i = q_i l_i, \quad i \in \{1, .., N\} \tag{3.2}$$

where l_i is the number of production workers employed and q_i is firm-specific technology. For simplicity, firms are assumed to have the same technology, $q_i = q \in \mathbb{R}_{++}$, $\forall i \in I$. Let w denote the wage rate. From (3.2) the marginal cost becomes w/q_i. Finally, the labor supply is assumed to be equal to 1. In this static version of the model, there is no incentive for conducting R&D since there is no future in which firms can harvest the return to their costly R&D investment. For existence of the equilibrium, it is enough to assume $AN > 2q\gamma$. In this environment, the equilibrium is characterized as follows.

The demand for each intermediate good is

$$y_i = \frac{A - 2Y p_i}{\gamma}, \quad i \in \{1, .., N\}.$$

Given this demand for its good, the monopolist maximizes its profit and chooses the following price and quantity:

$$y_i = \frac{Y}{\gamma} \left(\frac{A}{2Y} - \frac{w}{q_i} \right), \quad i \in \{1, .., N\}. \tag{3.3}$$

$$p_i = \frac{1}{2} \left(\frac{A}{2Y} + \frac{w}{q_i} \right), \quad i \in \{1, .., N\}. \tag{3.4}$$

The resulting profit function is

$$\pi_i = \frac{1}{2} \left(\frac{A}{2Y} - \frac{w}{q_i} \right)^2, \quad i \in \{1, .., N\}.$$

Finally, labor market clearing condition closes the model,

$$1 = \sum_{i=1}^{N} l_i \tag{3.5}$$

Given these specifications, the total number of unknowns of the model is $3N + 2 : Y, w,$ $[y_i]_{i=1}^{N}, [p_i]_{i=1}^{N}, [l_i]_{i=1}^{N}$. To solve for these unknowns we utilize equations $(3.1), (3.5)$ and N of

the equations (3.2), (3.3) and (3.4). The resulting equilibrium values are

$$l_i^* = \frac{1}{N}, \quad y_i^* = \frac{q}{N}$$
$$Y^* = \left(Aq - \frac{\gamma q^2}{2N}\right)^{1/2}$$
$$\frac{w^*}{Y^*} = \frac{AN - 2q\gamma}{2AN - q\gamma} < 1$$

3.2.2 Discussion

This simplified static version provides important insights on the intra-temporal interactions of the model.

Substitutability The first important observation is that both the final output and the labor share are increasing in the degree of substitution γ. The production function of the final good is concave in each of the intermediate good y_i and this curvature is governed by γ. As γ decreases, the intermediate goods become more substitutable and in the limit, $\gamma \to 0$, they become perfect substitutes. The degree of substitutability highlights the importance of the market size.

Market Size The second important observations is that both the final output and the labor share are increasing and the markups are decreasing in *market size* (N). Since the labor supply is fixed, an increase in market size implies a lower quantity produced per firm which leads to an increase in marginal product of each intermediate good. Therefore, everything else constant, a policy that increases the equilibrium number of firms in the economy is welfare improving.

Firm Size The third observation about the model is that the markups are increasing in technology q_i. In this model, firm size is monotonically related to technology therefore we can conclude that the markups are increasing in *firm size*.

Concavity Finally, except for a small range of values on the lower end, the profit function is concave in *firm's technology*. This feature will be crucial for the theoretical results. In the dynamic version of the model, when R&D is introduced, firm i will have incentives to improve its technology to increase its profit. However, these incentives will decrease as firm's technology q_i, or equivalently firm size y_i increases.

Welfare In this static model, there is no innovation. However in a multi-period setting,

116

a social planner who is willing to maximize the final output would target to improve the lowest technology in the economy. To see this, consider two types of firms with a measure of $N/2$ each: High productivity firms with q_h and low productivity firms with $q_l < q_h$. Then the profit maximizing output level is

$$y_i^* = \frac{1}{\gamma N} \left(\frac{AN}{2} - \frac{w}{q_i} \right), \quad i \in \{h, l\}$$

Once we substitute these values back into the production function we get

$$Y^* = \frac{A}{2\gamma} \left(AN - \frac{w}{q_h} - \frac{w}{q_l} \right) - \frac{1}{4\gamma N} \left[\left(\frac{AN}{2} - \frac{w}{q_h} \right)^2 + \left(\frac{AN}{2} - \frac{w}{q_l} \right)^2 \right]$$

Therefore the partial effects of the productivities have the following relation

$$\frac{\partial Y^*}{\partial q_l} = \frac{w}{4\gamma q_l^2} \left(A + \frac{2w}{Nq_l} \right) > \frac{w}{4\gamma q_h^2} \left(A + \frac{2w}{Nq_h} \right) = \frac{\partial Y^*}{\partial q_h}.$$

This exercise shows that if a policy, such as R&D subsidy, is going to be applied in order to increase output through improving the technologies of the firms, then this policy should target the small firms which are identified by q_l.

Transition to the Main Model   In the main model, the time horizon will be infinite and therefore firms will have incentives to invest in R&D and advance their technologies. I will relax the assumption that firms have identical technologies and use the R&D efforts to solve for the equilibrium invariant distribution of technologies.

In addition, I will endogenize the measure of firms, N and utilize free-entry condition to solve for it. This implies that the markups in the economy will be determined endogenously. Similarly, I will allow for exits and utilize free-exit condition to solve for the cut-off rule that the firms will adopt.

As a result of the R&D efforts, the technology aggregate $\sum_{i=1}^N q_i$ will grow at a rate g. The assumption of a positive externality of each innovation on TFP (A), will make sure that the economy has a balanced growth path on which the final output grows at the same rate, g.

The following section describes the main model in detail.

3.3 Main Model

3.3.1 Demographics, Preferences and Technology

Consider the following discrete time economy. The representative household maximizes the expectation of an infinite sum of discounted utility, with intertemporal preferences of the following form,

$$U_t = \mathbb{E}_t \sum_{\tau=t}^{\infty} \beta^\tau \ln C_\tau \tag{3.6}$$

where C_t denotes consumption at time t. In this specification, $\beta \in (0,1)$ is the discount factor and $\mathbb{E}_t$ the expectation operator conditional on the information at time t. The logarithmic form measures the per-period utility derived from the consumption of the final good Y_t which is produced using a continuum of differentiated goods indexed by $i \in I$. Specifically the production function takes the following form:

$$Y_t = \left[A_t \int_{i \in I} y_{i,t} di - \frac{\gamma}{2} \int_{i \in I} y_{i,t}^2 di - \frac{\delta}{2} \left(\int_{i \in I} y_{i,t} di \right)^2 \right]^{1/2} \tag{3.7}$$

where $y_{i,t}$ denotes the quantity of variety i and A_t is the economy-wide total factor productivity (TFP) at time t. Similar specifications have been used by several papers (Ottaviano et al (2002), Melitz and Ottaviano (2005), Corcos et al (2007)).[1] The parameters $\gamma, \delta > 0$ capture the substitution patterns among the varieties. The specification in (3.7) features *love for variety* which is measured by γ. Having $\gamma = 0$ implies that the varieties are perfect substitutes. A_t is assumed to grow proportionally to the technology aggregate (Q_t) through externalities of each individual innovation.[2] Without loss of generality, the price of the final good is set equal to 1. The household is allocated with a continuum of 1 unit of labor which will be supplied inelastically for production and R&D. The representative household also owns a balanced portfolio of all the firms in the economy which implies the following budget constraint

$$C_t \leq \int_{i \in I} \Pi_{i,t} di + w_t$$

where I is the set of active firms in the economy.

---

[1] I modify those previous specifications to make the production function compatible with a balanced growth path.

[2] The details about the technology aggregate Q_t is defined later in the model.

Final Good Sector

The final good producer maximizes profits, using a set of differentiated goods $[y_i]_{i \in I}$ as inputs and takes their prices as given. Each of the differentiated input $y_{i,t}$ is being produced by a monopolist $i \in I$ who charges $p_{i,t}$ for each unit it sells. While mapping the model into real life, one can think of each variety as a different brand of a product, say i and j as the Home Computing goods of Hewlett Packard (HP) and Dell. To capture the reality that HP produces and develops only HP products, in specification (3.7) firm i will compete with other firms from its own product line. In other words, it will only produce, improve and price products that are in its own product line i. The representative household demands all of the brands. In equilibrium, the final good producer will demand more of the lower priced products which will push firms to compete in lowering their costs of production.

The final good producer does not have any storage technology available and therefore optimizes period-by-period. The problem of the final good producer can be stated as

$$\pi_{Y,t} = \max_{[y_i]_{i \in I}} \left\{ Y_t \left(A_t, [y_i]_{i \in I} \right) - \int_{i \in I} p_{i,t} y_{i,t} di \right\} \tag{3.8}$$

This maximization generates the final good producer's demands for the intermediate goods,

$$y_{i,t}^d \left(p_{i,t} \right) = \max \left\{ 0, \frac{A_t - \delta D_t - 2 Y_t p_{i,t}}{\gamma} \right\}$$

where $D_t \equiv \int_I y_{i,t} di$ is the sum of the intermediate goods.

(Differentiated) Intermediate Goods Sector

This sector is composed of monopolistically competitive firms which make decisions about production, R&D, entry and exit. Each of those is described below.

Production   Monopolist i in the intermediate goods sector operates with the following linear technology,

$$y_{i,t} = q_{i,t} l_{i,t} \tag{3.9}$$

where $l_{i,t}$ is the labor hired by firm i for production at time t. Firm-specific labor productivity $q_{i,t} \in S_t \equiv \left[q_t^-, \infty \right)$ with a distribution function $\Psi_t \left(S_t \right)$ and $q_t^- \in \mathbb{R}_{++}$. Productivity $q_{i,t}$ can be stochastically improved through R&D. If we denote the wage rate by w_t, equation (3.9) implies

119

that the firm specific marginal cost of production is simply $c_{i,t} = w_t/q_{i,t}$. For convenience, I will define the aggregate technology index as $Q_t \equiv \int_{i \in I_t} q_{i,t} di$.

On the production side, monopolist i decides how much to produce and what price to charge for its product, taking the final good (Y_t), wage rate (w_t) and the productivity distribution (Ψ_t) in the economy as given. Its production decision is simply

$$\max_{p_{i,t}, y_{i,t}} \left\{ (p_{i,t} - c_{i,t}) y_{i,t} \mid Y_t, A_t, w_t, \Psi_t \right\} \tag{3.10}$$

$$subject\ to$$

$$y_{i,t} = y_{i,t}^d (p_{i,t})$$

R&D The evolution of the firm specific marginal cost at time t is determined by two factors: the growth rate of the wage rate and the innovation by the firm. Each intermediate good i has a quality ladder along which firm i improves its state of technology $q_{i,t}$ through *additive* step-by-step innovations. The outcome of R&D is uncertain, so that a firm cannot directly choose to make an innovation, but only a probability of success for achieving an innovation. The novelty of the current model is that firms can endogenously choose not only this probability of success for but also the quality of innovation. Hence, each R&D decision is splitted into two components:

- the quality of the innovation that the firm aims for, $\lambda_{i,t} \in \left[0, \bar{\lambda}_t\right]$, and

- the probability of success, $\phi_{i,t} \in [0,1]$.

Let $q_{i,t}$ be the technology of firm i at time t. Any successful innovation of quality $\lambda_{i,t}$ improves the technology by this amount with probability $\phi_{i,t}$. As the overall technology in the economy improves, the maximum step size (innovation quality) $\bar{\lambda}_t$ will grow as well. For simplicity, this upper limit will be set as $\bar{\lambda}_t = w_t$.[3] Let $\left(\lambda_{i,t}, \phi_{i,t}\right)$ be firm i's R&D decision in

---

[3] Since the payoff relevant state variable is q/w the payoff relevant improvement will become λ/w. Setting the upper limit as $\bar{\lambda} = w$ bounds the payoff relevant innovation size between 0 and 1, i.e., $\lambda/w \in [0,1]$.

period t. Consequently, the state of technology for firm i in the subsequent period will be

$$q_{i,t+1} = \begin{cases} q_{i,t} + \lambda_{i,t} \text{ with probability } \phi_{i,t} \\ \\ q_{i,t} \text{ with probability } 1 - \phi_{i,t} \end{cases} \quad (3.11)$$

The innovation step $\lambda_{i,t}$ is the size of the improvement on top of the current technology. This specification assumes that there are constant returns to scale to technology improvements; in other words, the improvements are independent of the current technology of the firm. The alternative specification, which has been used by several studies in the literature (Aghion, Harris and Vickers, 1997; Aghion, Howitt, Harris and Vickers, 2001; Acemoglu, 2008) is that the improvements are proportional to the current technology, yet this would introduce increasing returns to scale to technology improvements.[4]

The theoretical results of the current model will rely on two features: First, the concavity of the value function in technology and second, the absence of any strong increasing returns in technology. These two features would imply that firms with lower technology, will have a greater incentive to increase their productivity. Therefore, they are going to choose higher quality innovations and grow faster. The assumptions on the production function in (3.7) and the structure of the technology improvement in (3.11) guarantee those two aforementioned features.

The specification for the R&D cost function is as follows. Let

$$h_{i,t} = h\left(\frac{\lambda_{i,t}}{w_t}, \phi_{i,t} \right) \quad (3.12)$$

denote the amount of labor required to undertake an R&D project of size λ_i with a success probability of ϕ_i at time t. In real life, firms benefit from the improvement of the aggregate technology through (i) spillovers (Jaffe, 1989; Jaffe, Trajtenberg, and Henderson, 1993; Anselin, Varga, and Acs, 1997, 2000) (ii) improvements in labor substituting capital in R&D (computers, for instance). The reduced form (3.12) captures such positive externalities. In steady-state, w_t is a fraction of Y_t, therefore, (3.12) implies that a given quality of innovation requires less labor as

---

[4]One way to eliminate this increasing returns to scale would be to make the R&D cost function proportional to the current quality.

the overall technology in the economy advances. The function h is assumed to be strictly convex, with $h_x(\cdot,\cdot), h_{xx}(\cdot,\cdot) > 0$, $h_{xz}(\cdot,\cdot) \geq 0$; $\frac{h_x(\cdot,\cdot)}{z}$ decreases in z, for $x, z \in \{\lambda, \phi\}$ and $x \neq z$, and the Inada conditions are assumed to hold $h(0, \cdot), h(\cdot, 0) = 0$, $h(1, \cdot), h(\cdot, 1) = \infty$. As a result, the cost of an R&D project $(\lambda_{i,t}, \phi_{i,t})$ is simply the R&D labor expense, $w_t h_{i,t}$. Having labor as the only input for both R&D and production and the specification for the supply of the numeraire good imply the following resource constraint for the economy,

$$C_t \leq Y_t.$$

Next I assume that each innovation has a positive externality on the aggregate TFP A. Therefore, in equilibrium, TFP grows proportionally to the technology index Q. This assumption assures that production function in (3.7) is constant returns to scale, which is crucial for the balanced growth path.

Exit All firms are assumed to have an outside option of $\Lambda_t = \nu_\Lambda Y_t$ where $\nu_\Lambda \in (0,1)$. This means when the market value of a firm falls below Λ_t, it would be optimal to exit and utilize the outside option. The decision to exit will be denoted by $\kappa_{i,t} = 1$ if the firm decides to exit and $\kappa_{i,t} = 0$, otherwise. In terms of timing of the model, firms decide to stay or exit after the stochastic R&D outcome is realized. In summary, the decision variables of an incumbent monopolist firm are i) output $y_{i,t}$, ii) price $p_{i,t}$, iii) innovation quality $\lambda_{i,t}$, iv) innovation intensity $\phi_{i,t}$, and v) exit $\kappa_{i,t}$.

In addition to the voluntary exits, firms receive idiosyncratic destructive shocks with an exogenous probability ρ upon which firms exit the market.[5]

Entry The market has an outside pool of potential entrants, which is large enough to have entry as long as it is profitable. These outside firms are ready to pay an entrance fee $\chi_t = \nu_\chi Y_t$ where $\nu_\chi \in (0,1)$ and determine their entry level productivity $q_{i,t}$ with a draw from a uniform distribution $q \sim U_t \equiv U[q_{\min,t-1}, \bar{q}_{t-1}]$ where $q_{\min,t-1}$ and $\bar{q}_{t-1}$ are the lower bound and the mean of previous period's productivity distribution, $\Psi_{t-1}(q)$. In the beginning of period t, these potential entrants pay the fee and draw a productivity. As a result of this draw $q_{i,t} \sim U_t$,
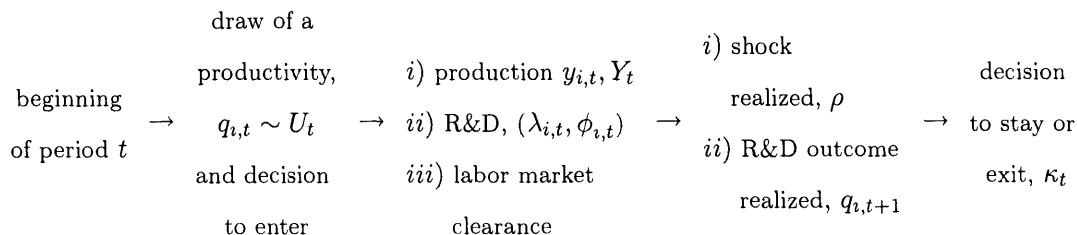
---

[5]One can also interpret this parameter as the degree of patent protection in the economy. If $\rho = 1$, there is no protection at all whereas $\rho = 0$ implies full protection.

firms with a high productivity draw will enter the market and the firms with a low productivity draw will find it more profitable to stay out of the market and utilize the outside option. The outside firms will enter the market as long as the expected value of entry is greater than the entry fee χ_t. This process will determine the equilibrium measure of the firms in the market, N_t.

Labor Market Labor is being employed by monopolist i both in production $(l_{i,t})$ and R&D $(h_{i,t})$. The measure of the labor supply is assumed to be 1. Given a wage rate w_t, labor market has to satisfy the following condition in equilibrium,

$$1 = \int_{i \in I} \left[l_{i,t} \left(w_t \right) + h_{i,t} \left(w_t \right) \right] di \tag{3.13}$$

Finally, to review the model, the timeline of the model in period t can be summarized as follows:

$$
\begin{array}{ccccccc}
& & \text{draw of a} & & & i) \text{ shock} & \\
& & \text{productivity,} & i) \text{ production } y_{i,t}, Y_t & & & \text{decision} \\
\text{beginning} & & & & & \text{realized, } \rho & \\
& \rightarrow & q_{i,t} \sim U_t & \rightarrow & ii) \text{ R\&D, } (\lambda_{i,t}, \phi_{i,t}) & \rightarrow & \rightarrow \text{ to stay or} \\
\text{of period } t & & & & & ii) \text{ R\&D outcome} & \\
& & \text{and decision} & iii) \text{ labor market} & & & \text{exit, } \kappa_t \\
& & & & & \text{realized, } q_{i,t+1} & \\
& & \text{to enter} & \text{clearance} & & &
\end{array}
$$

3.3.2 Equilibrium

Throughout the model, I will focus on the Markov Perfect Equilibrium which makes the payoffs functions of the payoff relevant state variables. For each individual firm, the state variables are the firm specific productivity $q_{i,t}$ and the productivity distribution Ψ_t in the economy. In the maximization problem of each monopolist, the wage rate is a sufficient statistics for the productivity distribution, therefore the only pay-off relevant state variable becomes the ratio of the firm's productivity to the wage rate, $\hat{q}_{i,t} \equiv q_{i,t}/w_t$. This conclusion allows me to drop firm specific indices and focus on the variables only as a function of the state space $\hat{q}_t$.

Before starting to characterize the equilibrium, I provide the definition of an allocation in this economy. Henceforth, bold letters will denote the entire mapping of the variables from the state space, i.e., $\mathbf{y}_t \equiv [y_{\hat{q},t}]_{\hat{q} \in \mathbb{R}_{++}}$.

123

Definition 3.1 (Allocation) *An allocation in this economy consists of i) consumption levels* $[C_t]_{t=0}^{\infty}$, *ii) prices, quantities, innovation qualities, intensities and exit decisions of monopolists* $[\mathbf{p}_t, \mathbf{y}_t, \boldsymbol{\lambda}_t, \boldsymbol{\phi}_t, \boldsymbol{\kappa}_t]_{t=0}^{\infty}$, *iii) the measure of available product types,* $[N_t]_{t=0}^{\infty}$, *iv) the distribution of productivities,* $[\boldsymbol{\Psi}_t]_{t=0}^{\infty}$ *and v) wage rates,* $[w_t]_{t=0}^{\infty}$.

Markov Perfect Equilibrium strategies can simply be represented as

$$[\mathbf{p}_t, \mathbf{y}_t, \boldsymbol{\lambda}_t, \boldsymbol{\phi}_t, \boldsymbol{\kappa}_t]^T : \mathbb{R}_{++}^3 \to \mathbb{R}_+^2 \times [0,1]^2 \times \{0,1\}$$

where the state-dependent strategies take the relative technology q_t/w_t, final output Y_t, and the wage rate w_t and map them to price, output, R&D decisions and exit decisions. The formal definition of the Markov Perfect Equilibrium goes as follows.

Definition 3.2 (Equilibrium) *A Markov Perfect Equilibrium is an allocation* $\langle [C_t^*]_{t=0}^{\infty}, [\mathbf{p}_t^*, \mathbf{y}_t^*, \boldsymbol{\lambda}_t^*, \boldsymbol{\phi}_t^*, \boldsymbol{\kappa}_t^*]_{t=0}^{\infty}, [N_t^*]_{t=0}^{\infty}, [w_t^*]_{t=0}^{\infty}, [\boldsymbol{\Psi}_t^*]_{t=0}^{\infty} \rangle$ *such that i)* $[\mathbf{p}_t^*, \mathbf{y}_t^*]_{t=0}^{\infty}$ *solves monopolist's profit maximization, ii)* $[\boldsymbol{\lambda}_t^*, \boldsymbol{\phi}_t^*]_{t=0}^{\infty}$ *solves optimal R&D investment problem, iii)* $[\boldsymbol{\kappa}_t^*]_{t=0}^{\infty}$ *solves monopolist's exit problem, iv)* $[C_t^*]_{t=0}^{\infty}$ *is consistent with the household's optimization, v)* $[N_t^*]_{t=0}^{\infty}$ *makes the free-entry condition hold as an equality, vi) the evolution of the productivity distribution* $[\boldsymbol{\Psi}_t^*]_{t=0}^{\infty}$ *is consistent with the R&D, entry and exit decisions of the firms and vii)* $[w_t^*]_{t=0}^{\infty}$ *is consistent with the labor market clearing condition.*

Having provided the definition of the equilibrium for this economy, I start to solve for it with the production decision of the monopolist.

Production The monopolist i with a marginal cost $c_{i,t} = w_t/q_{i,t}$ takes this inverse demand for its variety as given and maximizes its profit as in (3.10). Since the pay-off relevant state variable is the ratio of the technology to the wage rate, I will adopt the following notation $\hat{q}_{i,t} \equiv q_{i,t}/w_t = 1/c_{i,t}$. The maximization of the monopolist delivers the following output and price decisions

$$y_{i,t} = \begin{cases} \bar{y}_t + \frac{Y_t}{\gamma}\left[\overline{\left(\frac{1}{\hat{q}_t}\right)} - \frac{1}{\hat{q}_{i,t}}\right] & \text{for } \hat{q}_{i,t} > \underline{\hat{q}_t} \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

$$p_{i,t} = \frac{1}{2}\left[\frac{\gamma \bar{y}_t}{Y_t} + \overline{\left(\frac{1}{\hat{q}_t}\right)} + \frac{1}{\hat{q}_{i,t}}\right] \tag{3.15}$$

where $\underline{\hat{q}_t} = \left[\frac{\gamma \bar{y}_t}{Y} + \overline{\left(\frac{1}{\hat{q}_t}\right)}\right]^{-1}$ and $\overline{\left(\frac{1}{\hat{q}_t}\right)} \equiv \frac{1}{N}\int\left(\frac{1}{\hat{q}_t}\right)di$. The optimal price set by the monopolist increases in its marginal cost $c_{i,t} = 1/\hat{q}_{i,t}$. Note that with the production function in (3.7)

the mark-ups are decreasing in the number of firms and in marginal cost. This implies that an increase in marginal cost is always shared by both the monopolist and consumer. The empirical findings of Hopenhayn and Campbell (2002) and Syverson (2004, 2007) empirically support these features.

Next, using the optimal pricing and output decisions of the monopolists, its profit becomes the following expression

$$
\pi_t\left(\hat{q}_{i,t}\right) = \begin{cases} \frac{Y}{2\gamma}\left(\frac{\gamma\bar{y}_t}{Y} + \overline{\left(\frac{1}{\hat{q}_t}\right)} - \frac{1}{\hat{q}_{i,t}}\right)^2 & \text{for } \hat{q}_{i,t} > \underline{\hat{q}_t} \\ 0 & \text{otherwise} \end{cases} \tag{3.16}
$$

where $\overline{\left(1/\hat{q}\right)}$ is the average marginal cost in the economy. Profits decrease in marginal costs, providing firms with an incentive to reduce their costs through innovations. Another key point regarding (3.16) is that it is concave for $\hat{q}_{i,t} > 3\underline{\hat{q}_t}/2$ which will be a key feature for the results.

Value Functions, R&D and Exit   Starting from the end of period t, we can formulate the maximization of firm i backwards. Let $q_{i,t}$ be the production technology of firm i in period t and let $q_{i,t+1}$ be the technology after the R&D outcome is realized at the end of period t. Denoting the beginning-of-period and end-of-period values of firm i by $V_t\left(\cdot\right)$ and $W_t\left(\cdot\right)$, respectively, and recalling that the pay-off relevant state variable is $\hat{q}_t \equiv q_t/w_t$, the exit problem can be stated as

$$
W_t\left(\hat{q}_{i,t+1}\right) = \max_{\kappa_i\in[0,1]}\left\{\kappa_{i,t}\Lambda_t + \left(1 - \kappa_{i,t}\right)V_{t+1}\left(\frac{\hat{q}_{i,t+1}}{1 + g_t}\right)\right\}. \tag{3.17}
$$

where Λ_t is the outside option and g_t is the growth rate of the wage rate between time t and $t+1$ which firm i takes it as given. Now going backwards, conditional on the fact that the firm starts the period with a productivity level of $q_{i,t}$, the program for the R&D decision is summarized by the following Bellman equation,

$$
V_t\left(\hat{q}_{i,t}\right) = \max_{\substack{\lambda_i\in[0,\bar{\lambda}_t]\\ \phi_i\in[0,1]}}\left\{\begin{array}{c} \pi_t\left(\hat{q}_{i,t}\right) - w_t h\left(\hat{\lambda}_{i,t}, \phi_{i,t}\right) + \frac{\rho\Lambda_t}{1+r} \\[2mm] \frac{1-\rho}{1+r}\left[\phi_{i,t}W_t\left(\hat{q}_{i,t} + \hat{\lambda}_{i,t}\right) + \left(1 - \phi_{i,t}\right)W_t\left(\hat{q}_{i,t}\right)\right] \end{array}\right\} \tag{3.18}
$$

where $\hat{q} \equiv q/w$, $\hat{\lambda} \equiv \lambda_t/w_t$, ρ and r are the exogenous destruction probability and the interest rate, respectively. The first term in this equation is the gross profit, the second expression is

125

the R&D expenditure due to the hired R&D workers for an innovation project with a quality $\lambda_{i,t}$ and a success probability $\phi_{i,t}$. The third term is the discounted value of being hit by the exogenous negative shock with probability ρ. The expression in the bracket is the expected end-of-period value of firm i taking into account that the R&D investment will be successful with the endogenous probability $\phi_{i,t}$.

Entry In the beginning of period t, potential entrants pay an entrance fee $\chi_t = \nu_\chi Y_t$ and determine their starting technology with a draw from the uniform distribution, $\hat{q} \sim U_t \equiv U\left[\hat{q}_{\min,t-1}, \overline{\hat{q}}_{t-1}\right]$. Firms will attempt to enter, as long as the expected value of entry is higher than the entry fee. This *free-entry condition* reads as,

$$\int V_t(\hat{q}) \, dU_t \geq \chi_t. \tag{3.19}$$

When the expected value of entry is greater than the entry fee, the measure of firms in the market, N_t will implicitly increase leading to a uniform decrease in the value function since the profits are decreasing in the number of firms. As a result, the expectation on the left-hand side of (3.19) will decrease until the inequality holds with equality.

Labor Market Labor is being employed both in the production and in the R&D sector. From equation (3.14), the demand for production worker is

$$l_{i,t} = \begin{cases} \frac{1}{w_t} \left(\frac{\bar{y}_t}{\hat{q}_{i,t}} + \frac{Y_t}{\gamma \hat{q}_{i,t}} \left[\overline{\left(\frac{1}{\hat{q}_t}\right)} - \frac{1}{\hat{q}_{i,t}} \right] \right) & \text{for } \hat{q}_{i,t} > \underline{\hat{q}}_t \\ 0 & \text{otherwise} \end{cases}$$

Previous section showed that firm i's labor demand for R&D activity is $h(\hat{q}_{i,t})$. Setting the supply of labor in the economy to 1, the labor market clearing condition is

$$1 = \frac{N_t}{\tilde{w}_t} \left(\frac{\bar{y}_t}{Y_t} \overline{\left(\frac{1}{\hat{q}_t}\right)} + \frac{1}{\gamma} \left[\overline{\left(\frac{1}{\hat{q}_t}\right)}^2 - \overline{\left(\frac{1}{\hat{q}_t^2}\right)} \right] \right) + N_t \bar{h}_t$$

where $\bar{y}$, $\overline{(1/\hat{q}_t)}$ and $\bar{h}$ stand for the average values of differentiated goods, normalized marginal costs and the workers employed by firms for R&D, respectively. From this equation, it is clear that the relevant variable for the labor market clearing is the labor share, $\tilde{w}_t = w_t/Y_t$. If the demand for labor is lower than the supply, the normalized wage rate $\tilde{w}$ would adjust to bring

the market into equilibrium.

The focus, henceforth, will be on the steady-state, in which all aggregate variables grow at the same rate $g > 0$.

3.3.3 Steady-State and Theoretical Results

In order to make the problem stationary, let us start by normalizing all growing variables by Y_t and denote the normalized value of some generic variable x by $\tilde{x} \equiv x/Y$.

Definition 3.3 (Steady-State Equilibrium) *A Steady-State Equilibrium is a tuple* $\langle \mathbf{V}, \mathbf{W}, \boldsymbol{\lambda}^*, \boldsymbol{\phi}^*, \boldsymbol{\kappa}^*, \hat{\boldsymbol{\Psi}}^*, N^*, \tilde{w}^*, g^* \rangle$ *such that i)* $\mathbf{V}, \mathbf{W}$ *satisfy* (3.17) *and* (3.18), *ii)* $\boldsymbol{\lambda}^*, \boldsymbol{\phi}^*, \boldsymbol{\kappa}^*$ *solve the the value functions in* (3.17) *and* (3.18), *iii)* $\hat{\boldsymbol{\Psi}}^* (\hat{q})$ *forms an invariant distribution, iv)* $\tilde{w}^*$ *clears the labor market, v)* N^* *is constant and consistent with steady-state free-entry condition, vi) the aggregate variables* Y_t, w_t, *and the technology index* $Q^* = \int_I q_{i,t} di$ *grows at the steady-state rate* g^* *which is consistent with the steady-state R&D choices* $(\boldsymbol{\lambda}^*, \boldsymbol{\phi}^*)$.

Next I characterize the steady-state equilibrium. The normalized steady-state versions of the value functions in (3.18) take the following form

$$\tilde{V}(\hat{q}_i) = \max_{\substack{\lambda_i \in [0,1] \\ \phi_i \in [0,1]}} \left\{ \begin{array}{c} \tilde{\pi}(\hat{q}_i) - \tilde{w}h\left(\hat{\lambda}_i, \phi_i\right) + \frac{\rho \tilde{\Lambda}}{1+r} \\ \\ \frac{1-\rho}{1+r}\left[\phi_i \tilde{W}\left(\hat{q}_i + \hat{\lambda}_i\right) + (1 - \phi_i)\tilde{W}(\hat{q}_i)\right] \end{array} \right\} \quad (3.20)$$

where $\hat{q} \equiv q/w$ and $\hat{\lambda} \equiv \lambda/w$ are the pay-off relevant state variable and normalized innovation quality. Similarly the end-of period value function becomes

$$\tilde{W}(\hat{q}_i) = \max_{\kappa_i \in [0,1]} \left\{ \kappa_i \tilde{\Lambda} + (1 - \kappa_i)\tilde{V}\left(\frac{\hat{q}_i}{1+g^*}\right) \right\}. \quad (3.21)$$

Since the transformed problem is autonomous, time subscripts are dropped for notational convenience. A point worth noting is that the growth of the wage rate g reflects the pressure of the rest of the firms on firm i. The marginal cost of firm i increases as the rest of the firms innovate because the increase in aggregate technology reflects itself through an increase in the wage rate. This implies that for any given $g > 0$, if firm i never innovates, it will have to eventually exit the market. Therefore, in addition to the usual cost reduction incentive, there is a supplementary incentive to innovate, namely to survive in the market.

The optimal steady-state exit decision in (3.21) is

$$\kappa_i^* = \begin{cases} 1 \text{ if } \tilde{\Lambda} > \tilde{V}\left(\frac{\hat{q}_{i,t+1}}{1+g^*}\right) \\ 0 \text{ if } \tilde{\Lambda} < \tilde{V}\left(\frac{\hat{q}_{i,t+1}}{1+g^*}\right) \end{cases}$$
$$\kappa_i^* \in [0,1] \text{ otherwise}$$

This simple condition implies that if the firm value is below the normalized outside option $\tilde{\Lambda}$, the firm will find it optimal to exit.

The next proposition characterizes the value function in (3.20). Its proof will be presented in Appendix B of the chapter, together with all other proofs.

Proposition 3.1 *Consider the dynamic optimization of an individual firm. For any given steady-state values of* $\left(\hat{\Psi}^*, N^*, \tilde{w}^*, g^*\right)$

i) *the value function in (3.20) exists, is unique, continuous and increasing,*

ii) *optimal policies* $\lambda^*\left(\hat{q}\right), \phi^*\left(\hat{q}\right)$ *exist, are compact-valued and upper hemi-continuous,*

iii) *for* $\hat{q} > \frac{3\hat{q}(1+g^*)}{2}$, *the value function in (3.20) is differentiable and strictly concave.*

This proposition documents that the firms with better technology have higher steady-state value. The following Lemma follows immediately from Proposition 3.1.

Lemma 3.1 *Let* $\hat{q}_{\min}$ *be defined implicitly by* $\tilde{\Lambda} = \tilde{V}\left(\hat{q}_{\min}/\left(1+g^*\right)\right)$. *Then firm i's exit decision is a cut-off rule such that*

$$\kappa_i^* = 1 \text{ if } \hat{q}_i < \hat{q}_{\min}$$
$$\kappa_i^* = 0 \text{ if } \hat{q}_i > \hat{q}_{\min}$$
$$\kappa_i^* \in [0,1] \text{ otherwise}$$

Recall that the marginal cost w/q_i increases if the firm fails to innovate since $w_t = w_0\left(1+g^*\right)^t$ grows at the rate g^*. Together with the cut-off rule, this implies that the firms below a certain threshold $\hat{q}_i < \left(1+g^*\right)\hat{q}_{\min}$ will exit the market with certainty if they fail to innovate in the current period. I will call the active firms below this threshold, $\hat{q}_i \in [\hat{q}_{\min}, \hat{q}_{\min}\left(1+g^*\right))$ *distressed* firms.

Next, having demonstrated the differentiability and concavity of the value function in Proposition 3.1, we can use the first order conditions to pin down the optimal steady-state R&D decisions for $\hat{q} > \frac{3\hat{q}(1+g^*)}{2}$:

$$\hat{\lambda}^{*}(\hat{q}) \quad : \quad \frac{(1 + g^{*})\,\tilde{w}^{*}h_{\hat{\lambda}}\left(\hat{\lambda}^{*}, \phi^{*}\right)(1 + r)}{\phi^{*}} = \tilde{V}'\left(\frac{\hat{q} + \hat{\lambda}^{*}}{1 + g^{*}}\right) \tag{3.22}$$

$$\phi^{*}(\hat{q}) \quad : \quad \tilde{w}^{*}h_{\phi_{i}}\left(\hat{\lambda}^{*}, \phi^{*}\right)(1 + r) = \tilde{V}\left(\frac{\hat{q} + \hat{\lambda}^{*}}{1 + g^{*}}\right) - \tilde{V}\left(\frac{\hat{q}^{*}}{1 + g^{*}}\right) \tag{3.23}$$

The key economic force for the innovation size is the marginal value of the new state that firm i is going to reach if successful. On the other hand, the incentive for the innovation intensity is the difference between the values of the successful state and the failure state, which is simply the private value of innovation. The larger this private value is, the more intensively firm i will try to innovate.

Let $(\boldsymbol{\lambda}^{*}, \boldsymbol{\phi}^{*})$ denote the steady-state R&D decisions of all the firms in (3.18). Then the aggregate steady-state growth rate is

$$g^{*} = \frac{(1 - \rho)\int_{I^{*}}\left(\hat{\lambda}_{i}^{*}\phi_{i}^{*} + \hat{q}_{i}\right)di}{\int_{I^{*}}\hat{q}_{i}\,di - \rho\left(\hat{q}_{\min} + \bar{\hat{q}}\right)/2} - 1 \tag{3.24}$$

where ρ is the probability of a destructive shock, $\hat{q}_{\min}$ and $\bar{\hat{q}}$ are the minimum and average normalized productivities.[6] This expression makes it clear that the aggregate growth rate is determined not only by the heterogenous innovation intensities ϕ^{*}, but also by the heterogenous innovation qualities, $\boldsymbol{\lambda}^{*}$. This observation has an important implication: Any policy instrument, such as R&D Subsidy, R&D Tax Credit or Intellectual Property Rights Policies, which targets to improve the aggregate economic growth through technological innovation should also take into account the heterogeneity of the innovation qualities of different firms. The following theorem states the prediction of the model about the relation between innovation quality and firm size. It is consistent with the third reduced form evidence from chapter 2 which stated that smaller firms produce higher quality innovations.

Theorem 3.1 (Innovation Quality) *Let* $\hat{\lambda}(\hat{q}) : \mathbb{R}_{++} \to [0, 1]$ *be the policy function as described in Proposition* 3.1, *ii. For* $\hat{q} > \frac{3\hat{q}(1 + g^{*})}{2} \in \mathbb{R}_{++}$, $\hat{\lambda}(\hat{q})$ *is a monotonically decreasing function such that* $\lim_{\hat{q} \to \infty} \hat{\lambda}(\hat{q}) = 0$.

---

[6]See Appendix B for the derivation.

129

According to this result, for firms greater than a certain size threshold, innovations become less drastic as firm size increases. The intuition for this result comes from the shape of the value function. The shape of the value function is driven both by the concavity of the profit function in the production technology. Since the marginal value of innovation is diminishing, the incentives for drastic innovation decreases as firm size increases.

Given, this solution, the firm entry decisions can be discussed. After the firms with productivity $\hat{q}_i$ above $\hat{q}_{\min}$ enter the market, they invest in R&D, choose $\left(\hat{\lambda}\left(\hat{q}_i\right), \phi\left(\hat{q}_i\right)\right)$ and hire R&D workers accordingly. At any point in time, firms are assumed to have a normalized outside value of $\tilde{\Lambda} > 0$.

When the economy has a strictly positive growth rate $g^* > 0$, the marginal cost of firm i will increase in every period. If the firm's step size is bigger than the aggregate increase in productivity, i.e., if $\hat{\lambda}_i > g^*\hat{q}_i$, then the firm's relative technology will improve. From Theorem (3.1) we know that the quality of innovation, $\hat{\lambda}(\hat{q})$ is a strictly decreasing function converging to 0. Therefore, the following lemma is immediate.

Lemma 3.2 *Let $\hat{\lambda}^*(\hat{q})$ be the equilibrium optimal choice for the innovation quality. Then*

i) there exists $\hat{q}_{\max} \in \hat{S}$ such that $\hat{\lambda}^*(\hat{q}_i) \leq g^*\hat{q}_{\max}$, $\forall \hat{q}_i > \hat{q}_{\max}$, and

ii) for $\hat{q}_i > \hat{q}_{\max}$, $\hat{q}_{i,t+1} \leq \hat{q}_{i,t}$.

This lemma implies that there exists a threshold level of technology above which all states are *transient*. This will be crucial for the existence of the invariant distribution. The formal statement for the existence of the invariant distribution is provided in the proposition below.

Proposition 3.2 *Consider the above model and let the state of outside firms be denoted $\hat{q}_{out}$. For any given equilibrium R&D decisions in (3.22), (3.23) there exists a unique steady-state distribution of industries, $\hat{\Psi}^*(\hat{q})$ with an atom at $\hat{q}_{out}$. Moreover, the invariant distribution is continuous in its transition probabilities.*

Now we have all the necessary tools for the existence of the equilibrium, which corresponds to a nine-tuple $\left\langle N, \tilde{w}, g, \lambda(.), \phi(.), \kappa(.), \hat{\Psi}, \tilde{V}(.), \tilde{W}(.)\right\rangle$. The following lemma simplifies its construction.

Lemma 3.3 *Consider the economy described above and let $K \equiv \gamma\frac{\tilde{y}}{\tilde{Y}} + \bar{c}$. For the existence of an equilibrium, it is necessary and sufficient to show that equilibrium values of $\langle \tilde{w}, g, K \rangle$ exist.*

130

The intuition for this lemma comes from the fact that, as in expressions $(3.20) - (3.23)$, equilibrium firm decisions and value functions can be represented as the solution of a simple maximization problem where firms take the equilibrium values of $\langle N, \tilde{w}, g, K, \hat{q}_{\min} \rangle$ as given. Moreover, proposition 3.2 proves the existence of the resulting invariant distribution. Finally, the value function of each firm is increasing in $\hat{q}_{\min}$ and decreasing in N. Therefore any outside option $\tilde{\Lambda}$ and entry fee $\tilde{\chi}$ correspond to a compact-convex-set-valued $\hat{q}_{\min}$ and N. Hence, we can treat the values of $\hat{q}_{\min}$ and N as parameters of the firm's maximization problem and compute the corresponding values of $\tilde{\Lambda}$ and $\tilde{\chi}$ as a result of this maximization.

I establish the existence of the equilibrium by defining the problem as a fixed point problem and using Kakutani's Theorem. The idea of the proof is that the equilibrium is the fixed point of a upper hemi-continuous operator that maps elements from a three dimensional compact-convex set *into* the same set. Let $m = \langle \tilde{w}, g, K \rangle \in \mathbf{M}$ be a generic argument of the following operator

$$\Phi(m) : \mathbf{M} \to \mathbf{M} \tag{3.25}$$

such that

$$\Phi(m) \equiv \begin{bmatrix} \Phi_{\tilde{w}}(m) \\ \Phi_g(m) \\ \Phi_\kappa(m) \end{bmatrix}$$

It is shown in the Appendix that all of the individual operators in $\Phi(m)$ are upper hemi-continuous and $\mathbf{M}$ is a compact-convex set. These steps lead to the following proposition.

Proposition 3.3 *Consider the economy described above. A steady-state equilibrium* $m^* \equiv \left[N^*, \tilde{w}^*, g^*, \lambda^*(.), \phi^*(.), \kappa^*(.), \hat{\Psi}^*, \tilde{V}^*(.), \tilde{W}^*(.) \right]$ *of this economy exists. Moreover, the steady-state growth rate is strictly positive* $g^* > 0$.

Next, the second major result of the model, documenting the relationship between the R&D intensity and firm size can be stated:

Theorem 3.2 (R&D Intensity) *Let* $\mathfrak{R}(\hat{q}) : \mathbb{R}_{++} \to \mathbb{R}_+$ *be the R&D intensity function defined as*

$$\mathfrak{R}(\hat{q}) \equiv \frac{w^* h(\hat{q})}{p(\hat{q}) y(\hat{q})}.$$

For $\hat{q} > \frac{3\hat{q}(1+g^*)}{2} \in \mathbb{R}_{++}$, $\mathfrak{R}(\hat{q})$ *is a monotonically decreasing function.*

This result is in line with the second reduced form evidence from chapter 2. The intuition for this result is that the firm's incentives for radical innovations diminish as firm size increases. Firms turn to more incremental innovations as their size increases and this decreases their R&D intensity.

Firm Growth The engine of both firm level and aggregate growth in this economy is productivity enhancing innovation. The model delivered two main reasons for firms to engage in R&D, namely to increase profits and simply to survive in the market. As long as the improvement in the productivity dominates the increase in the wage rate, the marginal cost of production decreases and firms obtain a higher share in the market and grow. The following proposition characterizes the expected firm growth rate and documents the role that innovation quality plays for increasing firm growth.

Proposition 3.4 *Consider the model above and let $g^* > 0$ be the aggregate steady-state growth rate. Then, expected firm growth can be expressed as*

$$\mho(\hat{q}) = -\eta + (1-\eta)\left[g^* + (1+g^*)\frac{1 + \left(\frac{\phi_i \lambda_i}{q_i + \lambda_i} - 1\right)(1+g^*)}{K^* \hat{q}_i - 1}\right]$$

where $K \equiv \gamma\frac{\bar{y}}{Y} + \bar{c}$.

Intuitively, higher quality innovations bring higher growth, and since innovation quality decreases in firm size, it is expected that firm growth decreases in firm size, a result obtained in the reduced facts in chapter 2 and documented in the following theorem.

Theorem 3.3 (Firm Growth) *Let $\mho(\hat{q}) : \mathbb{R}_{++} \to \mathbb{R}$ be the function of expected growth rate as defined in Proposition 3.4. For $\hat{q} > \frac{3\hat{q}(1+g^*)}{2} \in \mathbb{R}_{++}$, $\mho(\hat{q})$ is a monotonically decreasing function.*

Hence, the model constructed predicts each of the real-world facts observed previously. It concludes that relatively big firms (firms above a certain size threshold) in the model behaved consistently with the data. Since the Compustat sample in chapter 2 consists of relatively large firms compared to the whole population in the US manufacturing sector, the predictions of the model matched the nature of the Compustat sample as well. The next objective of this chapter is to incorporate these facts into policy analysis. The following section takes a step in this direction and estimates the structural parameters of the model.

3.4 Structural Estimation

The theoretical model can be simulated with different parameter values, taken to the real data and used to simulate the effects of different macroeconomic policies. In this respect, a simulation-based estimation technique, like Simulated Method of Moments (SMM) is particularly useful because it enables us to focus on the relevant moments of the model and the data. In this section, I will first provide the necessary background information on SMM. Then I will describe the moments chosen and the computational strategy. Finally, I will conclude the section by presenting the estimates. The next section will conduct the policy experiment using these estimated parameters.

For the estimation, a parametric form of the R&D cost function needs to be specified and is chosen such as to satisfy all the assumptions previously imposed, namely:

$$h\left(\hat{\lambda}, \phi\right) = B_\lambda \frac{\hat{\lambda}^{\eta_\lambda}}{1 - \hat{\lambda}} + B_\phi \frac{\phi^{\eta_\phi}}{1 - \phi}$$

where $\eta_\lambda, \eta_\phi > 1$ and $B_\lambda, B_\phi > 0$. Separability is a natural benchmark assumption in this context since we do not have strong priors about the functional form of the R&D cost function with two arguments.

To specify the parameters to be estimated, the following simplifications are made: The discount rate is set at 95 percent per annum. The entry fee $\tilde{\chi}$ is also set a priori, such that the total measure of firms in the economy is equal to 1 because it is hard to find an informative moment condition for this parameter. The outside option ν_Λ determines the cut-off level for the exiters $\hat{q}_{\min}$. Alternatively, one can determine the cut-off level and compute the corresponding outside option. Since the mapping between the two is monotonic and continuous, this does not affect the estimates. For computational simplicity, I will follow the latter option. As a result, the vector of parameters of length $k = 8$ to be estimated, within the set Θ of feasible values is:

$$\theta = \begin{bmatrix} \gamma & \delta & B_\lambda & B_\phi & \eta_\lambda & \eta_\phi & \rho & \hat{q}_{\min} \end{bmatrix}^T \in \Theta$$

3.4.1 Simulated Method of Moments (SMM)

The rationale for using the SMM method lies in the lack of a closed form expression for the parameters in terms of the data moments. The idea of SMM is as follows. Let $\left(o_i^A\right)$, $i = 1, ..., n$

133

be i.i.d real observations and $\mathbb{M}^A$ denote the vector of the $l \geq k$ selected moments from the actual data. Let $\mathbb{M}^S(\boldsymbol{\theta})$ denote the vector of the corresponding moments that are generated from the simulation of the model for a given set of parameters $\boldsymbol{\theta}$. In addition, let $\mathbf{R}^2(\boldsymbol{\theta})$ denote the weighted sum of squared deviation (WSSD) between the data and their simulated counterparts,

$$\mathbf{R}^2(\boldsymbol{\theta}) = \left[\mathbb{M}^A - \mathbb{M}^S(\boldsymbol{\theta})\right]^T \mathbb{W} \left[\mathbb{M}^A - \mathbb{M}^S(\boldsymbol{\theta})\right] \tag{3.26}$$

where $\mathbb{W}$ is the optimal weighting matrix. SMM calculates the estimate $\hat{\boldsymbol{\theta}}$ by minimizing the distance between the data and the model moments,

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in \Theta} \mathbf{R}^2(\boldsymbol{\theta})$$

Due to possible discontinuities in the objective function, *Simulated Annealing Algorithm* is used for this minimization (See Goffe, 1996 for details of this algorithm).[7]

Let Ω denote the variance-covariance matrix of $(1/\sqrt{n})\left[\mathbb{M}^A - \mathbb{M}^S(\boldsymbol{\theta})\right]$. The optimal weighting matrix $\mathbb{W}$ is the inverse of the variance-covariance matrix $\mathbb{W} = \Omega^{-1}$ (Adda and Cooper, 2003, p. 88). Hence, observations with higher variance are given less weight. To calculate Ω, I follow Bloom (2009) and use a block bootstrap with replacement on the data.

Gouriéroux and Monfort (1996, p. 29) show that $\hat{\boldsymbol{\theta}}_n(\mathbb{W})$ is consistent when n tends to infinity and that

$$\sqrt{n}\left[\hat{\boldsymbol{\theta}}_n(\mathbb{W}) \to \boldsymbol{\theta}_0\right] \xrightarrow[n \to \infty]{d} N(0, Q(\mathbb{W}))$$

where

$$Q(\mathbb{W}) = \left[\frac{\partial \mathbb{M}^T}{\partial \theta} \mathbb{W} \frac{\partial \mathbb{M}}{\partial \theta^T}\right]^{-1}. \tag{3.27}$$

Therefore the magnitudes of the standard errors of estimates are determined both by the variance-covariance of the moments and the sensitivity of the moments to the parameters, with a greater sensitivity leading to smaller standard errors.

---

[7]It samples a $\boldsymbol{\theta} \in \Theta$ during each iteration and accepts the current $\boldsymbol{\theta}$ as the new candidate for the global minimizer with certainty if it reduces the WSSD and with some probability if the WSSD is increased with the current $\boldsymbol{\theta}$.

3.4.2 Data and Moments

In this estimation, I use the full Compustat manufacturing sample between 1980-2000. The identifying moment conditions are generated as follows: First, the means of the variables of interest are calculated. For each variable, the deviations from a linear time trend and from sector averages are taken. These deviations are added back to the corresponding variable means to obtain variables centered around their means, but purged from time and sector effects. After dropping outliers (the top and bottom 2.5 percentiles for each of the three variables firm growth, R&D intensity, firm sales), firms are ranked in terms of their size and divided into two bins of "small firms" and "large firms". Unfortunately, there is not a clear theory about how to pick the moment conditions. Therefore my strategy has been to include every possible informative moment condition in the data. The first fourteen moment conditions match the mean firm growth rates, R&D intensities and labor productivities, innovation qualities, labor shares defined as "the ratio of labor expense to sales", exit rates, patent count per R&D spending. The next moment condition matches the yearly transition rate from small firms to large firms. This way, 15 moment conditions are obtained to estimate 8 parameters.

3.4.3 Identification

The necessary condition for identification is that $\frac{\partial M}{\partial \theta^T}$ should have full column rank which means that the objective function $\mathbf{R}^2(\theta)$ has a unique local minimum attained for the true parameter value. For the efficiency of the estimator, it is essential to use informative moment conditions, that is the moments must be sensitive to changes in parameter values as shown in (3.27).

Hence, for identifying the cost parameters $\left(B_\lambda, B_\phi, \eta_\lambda, \eta_\phi\right)$, the moments used are the innovation quality, R&D intensity, innovation counts normalized by R&D spending and the transition rate which are directly determined by the R&D choices. To identify the exogenous exit rate (ρ), the exit rates in the data are used. To estimate the parameters of the production function in (3.6) and the cut-off productivity $(\gamma, \delta, \hat{q}_{\min})$, firm growth rates, labor shares and labor productivities are most useful.

135

3.4.4 Computational Strategy

The computational solution of the model consists of a nested fixed point problem. The outer layer of the nest consists of four variables, namely the aggregate growth rate g, the ratio of TFP to final output $\tilde{A} \equiv A/Y$, the sum of average marginal cost and normalized output $K \equiv \bar{c} + \gamma \bar{y}/Y$ and the labor share $\tilde{w}$. To solve the model, I start with some initial guess for these parameters. Next taken these values as given, inside the nest, firms' value functions are solved. The routine for solving the model is as follows.

1. Start with an initial argmax $\hat{\boldsymbol{\theta}}$ and set $\hat{\mathbf{R}}^2 = \infty$.

2. Sample a set of parameters $\boldsymbol{\theta}_{guess} \in \Theta$.

3. Start with a guess for $\left[g, \tilde{A}, K, \tilde{w} \right]_{guess}$

4. Solve the value functions

5. Generate the R&D efforts $[\lambda(\hat{q}), \phi(\hat{q})]_{\hat{q} \in \hat{Q}}$

6. Using $[\lambda(\hat{q}), \phi(\hat{q})]_{\hat{q} \in \hat{Q}}$ generate the invariant distribution of firms $\boldsymbol{\Psi}(\hat{q})$.

7. Solve for the new values of $\left[g, \tilde{A}, K, \tilde{w} \right]_{new}$

8. If $\left\| \left[g, \tilde{A}, K, \tilde{w} \right]_{guess}^T - \left[g, \tilde{A}, K, \tilde{w} \right]_{new}^T \right\| < criteria$, stop. Else, update $\left[g, \tilde{A}, K, \tilde{w} \right]_{guess}$ and go back to line 4.

9. Calculate the moments of the model $\mathbb{M}^S(\boldsymbol{\theta}_{guess})$. If $\mathbf{R}^2(\boldsymbol{\theta}_{guess}) < \hat{\mathbf{R}}^2$, then $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_{guess}$ and $\hat{\mathbf{R}}^2 = \mathbf{R}^2(\boldsymbol{\theta}_{guess})$. Else set $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_{guess}$ and $\hat{\mathbf{R}}^2 = \mathbf{R}^2(\boldsymbol{\theta}_{guess})$ with some probability which decreases as the number of iteration increases.

10. Repeat this loop as many times as possible. The more it is repeated, the more likely it will reach the global minimum.

3.4.5 Results

The following table reports the moments of the actual data and the model regarding the mean growth rate, R&D intensity, innovation quality, labor productivity, labor share, exit rates,

136

innovation counts normalized by R&D spending, and the transition rate from the small-firm

bin into the large-firm bin.

MOMENT CONDITIONS

Averages	Data		Model	
	Small Firms	Large Firms	Small Firms	Large Firms
Growth Rate	0.037	0.001	0.016	-0.001
R&D Intensity	0.082	0.040	0.099	0.047
Innovation Quality	3.892	3.365	3.871	3.175
Labor Productivity	1.330	1.695	1.409	1.674
Labor Share	0.286	0.278	0.388	0.317
Exit Rate	0.010	0.005	0.010	0.010
Innovation Count per R&D	0.014	0.026	0.028	0.038
Transition Rate	0.020		0.072	

TABLE 5

The average growth rate, R&D intensity and innovation quality are bigger among small firms.

More interestingly, the data shows that the labor share calculated as the ratio of labor expense

to sales is bigger among small firms. This is a natural conclusion of the theoretical model since

the mark-ups are increasing in firm size. The data also shows that the number of patents per

R&D spending is bigger among large firms and the labor productivity is higher among large

firms. All these qualitative facts have been matched by the simulated model. The following

table reports the parameter estimates and their standard errors.

PARAMETER ESTIMATES

	γ	δ	B_λ	B_ϕ	η_λ	η_ϕ	ρ	$\hat{q}_{\min}$
Estimate	1.430	3.850	0.683	0.312	2.440	1.743	0.010	5.090
St Dev	(.032)	(.009)	(.033)	(.017)	(.144)	(.138)	(.001)	(.157)

TABLE 6

Figure 1 shows the plots of the value function, output levels, innovation qualities and the firm size distribution at the estimated parameter values as a function of the firm technology. The value function is consistent with (3.1) and concave. Figure 1B shows the close link between firm size and its technology. Consistent with its theorem, figure 1C shows that innovation quality is decreasing in firm's technology. The next section utilizes these parameters in a policy experiment.

3.5 Policy Analysis: R&D Subsidy

The suboptimality of private R&D investment has spawned a heated debate both among academicians and policymakers. It is widely argued that the social return to R&D is greater than the private return, which suggests that the equilibrium amount of private R&D investment is suboptimal (Griliches, 1992, 1995; Jones and Williams, 1998). To align the social and private returns of R&D, policymakers in the U.S. and in the E.U. have used R&D subsidy programs. The UK, for instance, is one of the countries which employs size-dependent R&D policies,

reserving higher subsidies to SMEs[8] (Bloom et al 2001).

The implications of such R&D policies for the economy are substantial. Therefore it is essential to have a good framework to investigate their effects. This section will use the estimated theoretical model to analyse size-dependent R&D policies.

3.5.1 Theory

In a decentralized equilibrium, firms do not take into account the externalities and the monopoly distortions that they generate. In the current model, there are several externalities and distortions that differs the firm problem from social planner's problem. The first of these distortions is the monopoly markups that are increasing in firm size. Second, each innovation has a positive externality on the TFP A that is not internalized by the firm. Third, any given sized innovation requires less labor if the aggregate technology is more advanced, as in equation (3.12); therefore each innovation generates a positive externality through its contribution to the aggregate technology. Finally, each innovation shifts the productivity distribution so that the entrants draw a higher productivity on average over time.

As a result, the decentralized equilibrium does not match the first-best level. The R&D subsidy gives the policymaker a partial ability to align the private and the social incentives. In the following setup, the government provides a proportional R&D subsidy and finances $\tau_i \in [0, 1]$ portion of the whole R&D spending of the firm.[9] The subsidy rate τ_i is allowed to be size dependent and since firm size is a monotonic transformation of the state-variable $\hat{q}$, the

---

[8]SME: Small and Medium Sized Enterprises are defined as firms with less than 250 employees.
[9]Some policies considers subsidies for non-employee R&D spendings only. I will not focus on that.

tax scheme that I will consider is

$$\tau_i = \tau\left(\hat{q}_i\right), \quad \forall \hat{q}_i \in \hat{Q}$$

As a result, firm i will pay only $[1 - \tau\left(\hat{q}_i\right)] wh\left(\hat{q}_i\right)$ instead of the full R&D expenditure.

I will assume that the government finances these subsidies through lump sump corporate

taxes, $\mathcal{T} \geq 0$. With this policy, the value function of firm i becomes

$$\tilde{V}\left(\hat{q}_i\right) = \max_{\substack{\lambda_i \in [0,1] \\ \phi_i \in [0,1]}} \left\{ \begin{array}{l} \tilde{\pi}\left(\hat{q}_i\right) - \mathcal{T} - [1 - \tau\left(\hat{q}_i\right)] \tilde{w}h\left(\hat{\lambda}\left(\hat{q}_i\right), \phi\left(\hat{q}_i\right)\right) + \frac{\rho\tilde{\Lambda}}{1+r} \\ \\ \frac{1-\rho}{1+r}\left[\phi\left(\hat{q}_i\right) \tilde{W}\left(\hat{q}_i + \hat{\lambda}\left(\hat{q}_i\right)\right) + (1 - \phi\left(\hat{q}_i\right)) \tilde{W}\left(\hat{q}_i\right)\right] \end{array} \right\}.$$

The government follows a balanced budget. Therefore it must be the case that

$$\mathcal{T} = \int \tau\left(\hat{q}\right) wh\left(\hat{q}\right) d\Psi\left(\hat{q}\right)$$

In what follows, I will focus on optimal (welfare-maximizing) subsidy schedules.

Welfare In the steady state, the discounted sum in (3.6) can be expressed as

$$\begin{aligned} \text{Welfare}_0 &= \sum_{t=0}^{\infty} \beta^t \log\left(C_0\left(1 + g^*\right)^t\right) & (3.28) \\ &= \frac{\ln C_0}{1 - \beta} + g^* \sum_{t=0}^{\infty} \beta^t t \end{aligned}$$

140

where C_0 is the consumption at time 0. From the resource constraint of the economy we have

$$C_t = Y_t. \tag{3.29}$$

In order to compare the welfare levels of two steady-state economies, the only variable that needs to be specified exogenously is A which is the TFP in the final good sector. A is set to 1 and equation (3.28) is used to calculate welfare.

3.5.2 Alternative Regimes

Next, I use the estimated structural parameters of the model and introduce different policy regimes. As a benchmark, I first report the results when no R&D subsidy exists. In the current framework, this implies $\tau = 0$ and $\mathcal{T} = 0$. The following table reports the subsidy rate τ, lump sum tax $\mathcal{T}$, average innovation quality $\overline{\left(\hat{\lambda}\right)}$, average innovation probability $\bar{\phi}$, normalized wage rate or labor share $\tilde{w} \equiv w/Y$, initial consumption C_0, aggregate steady-state growth rate of the economy g^* and the resulting welfare in (3.28) :

No Subsidy

τ	$\mathcal{T}$	$\overline{\left(\hat{\lambda}\right)}$	$\bar{\phi}$	$\tilde{w}$	C_0	g^*	Welfare
0	0	0.255	0.271	0.068	0.301	0.010	-21.05

Table 7

Without any R&D subsidy, the average growth rate in the economy is 1% and the average innovation probability is 0.27. This implies that the expected duration of time between any two consecutive innovations is 4 years on average.

141

Uniform Subsidy

Next I introduce uniform R&D subsidy on the same group of firms, such that the subsidy rate does not depend on firm size. Formally, this corresponds to $\tau(\hat{q}) = \tau \in [0,1]$, $\forall \hat{q} \in \hat{Q}$. Under the optimal uniform subsidy policy, the model generates the following results:

OPTIMAL UNIFORM SUBSIDY

τ	$\mathcal{T}$	$\overline{\left(\hat{\lambda}\right)}$	$\bar{\phi}$	$\tilde{w}$	C_0	g^*	Welfare
0.711	0.068	0.469	0.533	0.207	0.263	0.092	8.377

TABLE 8

Without any subsidy, the growth rate of the economy was 1%. Subsidizing R&D made it cheaper for the firms to invest in R&D. As a result, the average innovation quality goes up from 0.25 to 0.47 and the expected duration goes down from 4 years to 2 years. The optimal uniform subsidy rate of 71% boosts growth to 9%, at the expense of a 13% lower initial consumption. In addition, the government collects a payment of 0.07 units worth of the final good as a lump-sum tax to finance these subsidies. Initial consumption is lower as scarce resources (labor in this model) are being diverted from production sector into R&D, but the overall effect on welfare is strongly positive, as welfare rises from -21.1 to 8.4 which is equivalent to 4-fold increase in the initial consumption from 0.3 to 1.2, had the growth rate remained the same at 1%.

142

Size-Dependent Linear Subsidy

The final policy regime that I consider is a linear subsidy policy, providing a differential subsidy

rate to firms of different sizes (productivities):

$$\tau\left(\hat{q}\right) = \beta_0 + \beta_1 \hat{q}, \quad \tau_1, \tau_2 \in \mathbb{R}, \quad \hat{q} \in \hat{Q}$$

The results are reported in the following table

<div align="center">

OPTIMAL LINEAR SUBSIDY

β_0	β_1	$\mathcal{T}$	$\left(\hat{\lambda}\right)$	$\bar{\phi}$	$\tilde{w}$	C_0	g^*	Welfare
1.597	-0.127	0.160	0.515	0.583	0.290	0.235	0.116	15.303

TABLE 9
</div>

Table 9 indicates that the optimal size-dependent subsidy rate is decreasing in firm size. Every

unit increase in productivity $\hat{q}$ reduces the subsidy rate by 0.127. This value implies that the

smallest firm in the distribution receives around 94% subsidy for its R&D investment whereas

the largest firm receives 58%. This policy generates an even bigger average innovation quality

of 0.52 and the increase in probability indicates that the expected duation between any two

consecutive innovation is reduced to 1.6 years. The growth rate increases to 11% and welfare

increases to 15.3 which is equivalent to having an initial consumption of 1.7 in the no-subsidy

case, a 6-fold increase.

The most interesting part of these results is the direction of the subsidy. The optimal policy

subsidizes small firms more heavily than large firms. The intuition for this result had been

provided in the simpler version of the model in section 3.2. Basically, the final good is concave

<div align="center">143</div>

in each of the technologies of firms, therefore the highest marginal product belongs to the firm with the lowest technology. The fastest way to increase the final good is to improve the lowest technology, hence the high subsidy goes to the firms with low technology which are also the small firms.

Overall, the results show that size-dependent R&D subsidy policies have a significantly positive impact on welfare. By providing higher subsidies to smaller firms, the private under-investment in R&D can be mitigated. Even though the cost of such policies tends to reduce consumption initially, the higher steady-state growth compensates for this and yields a higher welfare.

3.6 Conclusion

This chapter investigated the innovation dynamics of different sized firms and studied the relevant policy implications. It carried out its analysis in three distinct steps.

Firstly, a theoretical general equilibrium model was outlined, destined to explain the microeconomic causes of the reduced form relations in Chapter 2. In this model technologically heterogeneous firms compete for innovations to increase their operating profits. The major novelty of the model is that firms can endogenously choose both the probability of innovation and its quality. In line with the reduced form results, the key fact of the model is that smaller firms undertake more radical innovations. Furthermore the model also explains other stylized empirical patterns related to the relationship between firm size and firm growth and firm size and R&D intensity. Essentially, the concavity of the profit and the value functions, and the absence of strongly increasing returns to productivity provide greater incentives for smaller

144

firms to increase their productivity. They hence are more R&D intensive, choose higher quality innovations and grow faster.
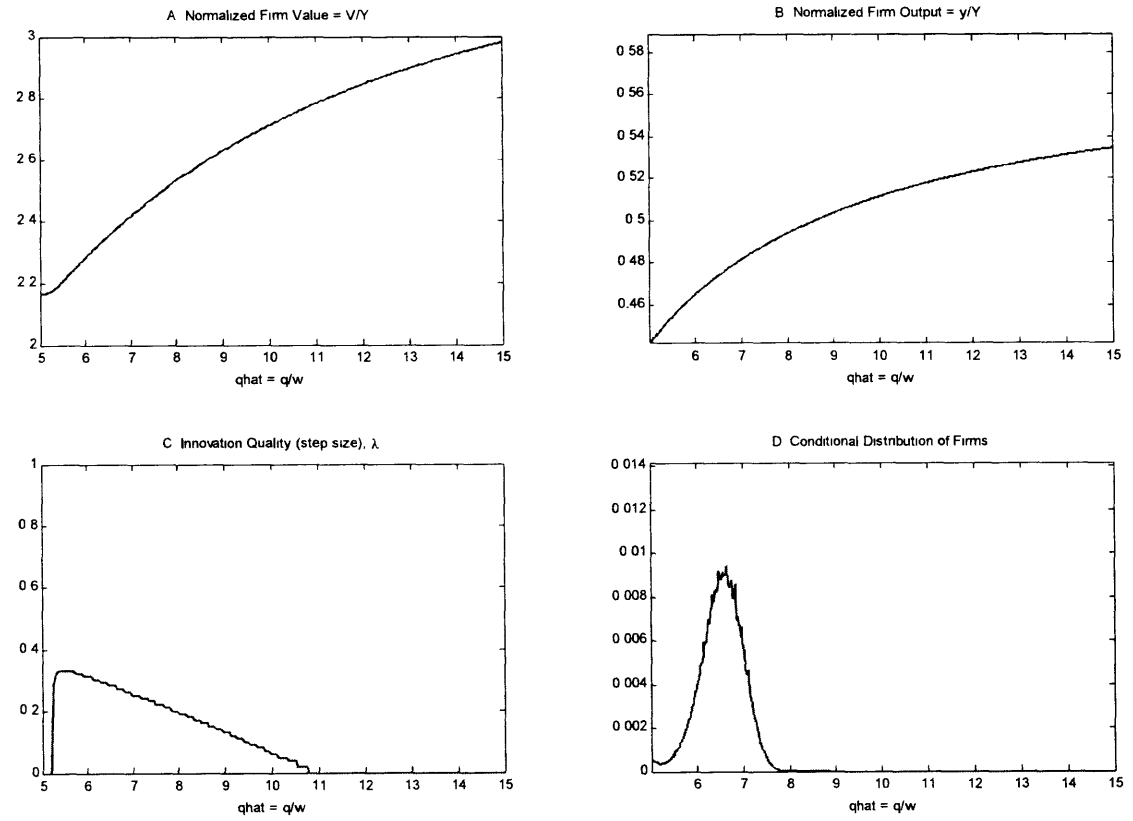
Second, the structural parameters of the model were estimated using the Simulated Method of Moments.

Third, these estimated parameters were then used for analyzing the effects of public R&D subsidies for different sized firms on innovation. Three different regimes were compared: *i)* No R&D subsidy, *ii)* a uniform (size-independent) R&D subsidy, *iii)* a size-dependent linear R&D subsidy. The results of this analysis documented significant gains from introducing R&D subsidies. Introducing uniform subsidy increased the growth rate from 1% to 9% and the welfare from -21.1 to 8.4 which was equivalent of 4-fold increase in the initial consumption. Next, moving from the uniform subsidy policy to the optimal size-dependent subsidy policy increased the growth rate to 12% and the welfare to 15 which corresponds to an additional 2-fold increase in the initial consumption. More interestingly, this welfare gain is achieved by providing higher R&D subsidy to small firms.

The results in this chapter documented the distinct and important contributions of small firms to aggregate technological innovations and growth. Size-dependent policies can allow policymakers to take advantage of this. A possible extension of the current work, which focuses exclusively on R&D subsidies, would be a study of the implications of entry and production subsidies on welfare. The challenge is to determine how close these alternative policies could bring the economy to the first best outcome.

3.7 Appendix A: Figure

FIGURE 1

A Normalized Firm Value = V/Y

B Normalized Firm Output = y/Y

C Innovation Quality (step size), λ

D Conditional Distribution of Firms

qhat = q/w

147

3.8 Appendix B: Proofs

Proof of Proposition (3.1).

$i - ii$) For any given values of $\left(\hat{\Psi}, N, \tilde{w}, g \right)$ the right-hand side of (3.20) defines a contraction mapping because it satisfies both *monotonicity* and *discounting* properties in Stokey, Lucas and Prescott, (1989). The per-period return function

$$\tilde{\pi}\left(\hat{q}_i \right) = \begin{cases} \frac{1}{2\gamma} \left(\frac{\gamma \bar{y}}{Y} + \overline{\left(\frac{1}{\hat{q}} \right)} - \frac{1}{\hat{q}_i} \right)^2 & \text{for } \hat{q}_i > \underline{\hat{q}} \equiv \frac{\gamma \bar{y}}{Y} + \overline{\left(\frac{1}{\hat{q}} \right)} \\ 0 & \text{otherwise} \end{cases}$$

is continuous in $\hat{q}$. In addition, it is bounded below by 0 and above by $\pi_{\max} = \frac{1}{2\gamma} \left(\frac{\gamma \bar{y}}{Y} + \overline{\left(\frac{1}{\hat{q}} \right)} \right)^2$. Moreover, $\left(\hat{\lambda}_i, \phi_i \right) \in \Gamma \equiv [0, 1]^2$ and Γ is clearly nonempty, compact-valued and continuous. Therefore the existence of the value function and the fact that the optimal policy is compact-valued and upper-hemicontinuous follow from Theorem 9.6 in Stokey, Lucas and Prescott (1989).

The per-period return function $\tilde{\pi}\left(\hat{q}_i \right)$ is increasing for $\hat{q}_i \in \left(\hat{q}_{\min}, \underline{\hat{q}} \right]$ and strictly increasing for $\hat{q}_i > \underline{\hat{q}}$. Besides Γ is increasing as well. Therefore the fact that the value function is (strictly) increasing follows from Theorem 9.7 in Stokey, Lucas and Prescott (1989).

iii) Since $\pi\left(\hat{q}_i \right)$ is strictly concave in $\hat{q}_i > 3\underline{\hat{q}}/2$, we have

$$\pi \left(\theta \hat{q}_i + (1 - \theta) \frac{\hat{q}_i + \hat{\lambda}_i}{1 + g} \right) > \theta \pi \left(\hat{q}_i \right) + (1 - \theta) \pi \left(\frac{\hat{q}_i + \hat{\lambda}_i}{1 + g} \right).$$

Moreover Γ is convex. Therefore the proof follows from Theorem 9.8 in Stokey, Lucas and Prescott (1989). ∎

Proof of Theorem (3.1). The proof follows from the fact that the value function is

increasing in $\hat{q}$. ∎

Proof of Theorem (3.1). For $\hat{q}_i > 3\hat{\underline{q}}/2$, the return function is strictly concave. From

Theorem 9.8 in Stokey, Lucas and Prescott (1989) it follows that the policy functions are single

valued and continuous. Let us recall the first order conditions,

$$\hat{\lambda}_i \quad : \quad \frac{\beta \phi_i}{1+g} v' \left(\frac{\hat{q}_i + \hat{\lambda}_i}{1+g} \right) = \tilde{w} h_{\hat{\lambda}} \left(\hat{\lambda}_i, \phi_i \right) \tag{3.30}$$

$$\phi_i \quad : \quad \beta \left[v \left(\frac{\hat{q}_i + \hat{\lambda}_i}{1+g} \right) - v \left(\frac{\hat{q}_i}{1+g} \right) \right] = \tilde{w} h_\phi \left(\hat{\lambda}_i, \phi_i \right) \tag{3.31}$$

We proceed case by case.

Case 1. Assume an initial equilibrium with $\left(\hat{q}_{old}, \hat{\lambda}_{old}, \phi_{old} \right)$ and by contradiction assume a new

equilibrium $\left(\hat{q}_{new}, \hat{\lambda}_{new}, \phi_{new} \right)$ such that $\hat{q}_{new} > \hat{q}_{old}$, $\hat{\lambda}_{new} \geq \hat{\lambda}_{old}$ and $\phi_{new} \leq \phi_{old}$. Then

(3.30) implies

$$\frac{\tilde{w} h_{\hat{\lambda}} \left(\hat{\lambda}_{new}, \phi_{new} \right)}{\phi_{new}} = \frac{\beta}{1+g} v' \left(\frac{\hat{q}_{new} + \hat{\lambda}_{new}}{1+g} \right)$$

$$< \frac{\beta}{1+g} v' \left(\frac{\hat{q}_{old} + \hat{\lambda}_{old}}{1+g} \right)$$

$$= \frac{\tilde{w} h_{\hat{\lambda}} \left(\hat{\lambda}_{old}, \phi_{old} \right)}{\phi_{old}}$$

$$\leq \frac{\tilde{w} h_{\hat{\lambda}} \left(\hat{\lambda}_{new}, \phi_{old} \right)}{\phi_{old}}$$

$$\leq \frac{\tilde{w} h_{\hat{\lambda}} \left(\hat{\lambda}_{new}, \phi_{new} \right)}{\phi_{new}}$$

which is a contradiction. Here the first line used the fact that v is concave (therefore its

slope is a decreasing function). The second line is just the initial equilibrium condition

and the last two lines are using the assumptions of the $h\left(.,.\right)$ function.

Case 2. Similarly assume an initial and new equilibria with $\hat{q}_{new} > \hat{q}_{old}$, $\hat{\lambda}_{new} \leq \hat{\lambda}_{old}$, $\phi_{new} \geq \phi_{old}$.

149

Define a function of

$$F\left(\hat{q}\right) \equiv \left[\frac{v\left(\frac{\hat{q}+\hat{\lambda}}{1+g}\right) - v\left(\frac{\hat{q}}{1+g}\right)}{\hat{\lambda}} \right].$$

This implies 2 subcases: **2.1)** $F\left(\hat{q}_{new}\right) \leq F\left(\hat{q}_{old}\right)$ and **2.2)** $F\left(\hat{q}_{new}\right) > F\left(\hat{q}_{old}\right)$.

Consider case 2.1. Then

$$
\begin{aligned}
F\left(\hat{q}_{new}\right) &\leq F\left(\hat{q}_{old}\right) \\
&= \frac{\tilde{w}h_\phi\left(\hat{\lambda}_{old}, \phi_{old}\right)}{\beta\hat{\lambda}_{old}} \\
&< \frac{\tilde{w}h_\phi\left(\hat{\lambda}_{old}, \phi_{new}\right)}{\beta\hat{\lambda}_{old}} \\
&\leq \frac{\tilde{w}h_\phi\left(\hat{\lambda}_{new}, \phi_{new}\right)}{\beta\hat{\lambda}_{new}}
\end{aligned}
$$

which is a contradiction.

Now consider subcase 2.2. Since $\hat{\lambda}$ and ϕ are continuous in $\hat{q}$, we might have 2 alternative cases again: $\exists \hat{q}^* > \hat{q}$ such that

Case 2.2.1 : $F\left(\hat{q}^*\right) = F\left(\hat{q}_{old}\right)$ and $\phi^* > \phi_{old}$ or **Case 2.2.2** : $F\left(\hat{q}^*\right) \geq F\left(\hat{q}_{old}\right)$ and $\phi^* = \phi_{old}$.

Consider case 2.2.1 :

$$
\begin{aligned}
F\left(\hat{q}^*\right) &= F\left(\hat{q}_{old}\right) \\
&= \frac{\tilde{w}h_\phi\left(\hat{\lambda}_{old}, \phi_{old}\right)}{\beta\hat{\lambda}_{old}} \\
&< \frac{\tilde{w}h_\phi\left(\hat{\lambda}_{old}, \phi^*\right)}{\beta\hat{\lambda}_{old}} \\
&\leq \frac{\tilde{w}h_\phi\left(\hat{\lambda}^*, \phi_{new}\right)}{\beta\hat{\lambda}^*}
\end{aligned}
$$

which is a contradiction to the fact that $\left(\hat{q}^*, \hat{\lambda}^*, \phi^*\right)$ is a solution.

150

Next consider case 2.2.2 : $F\left(\hat{q}^*\right) \geq F\left(\hat{q}_{old}\right)$ and $\phi^* = \phi_{old}$. This implies $\hat{q}^* + \hat{\lambda}^* < \hat{q}_{old} + \hat{\lambda}_{old}$.

Then

$$
\begin{aligned}
\frac{\beta\phi^*}{1+g}v'\left(\frac{\hat{q}^* + \hat{\lambda}^*}{1+g}\right) &> \frac{\beta\phi_{old}}{1+g}v'\left(\frac{\hat{q}_{old} + \hat{\lambda}_{old}}{1+g}\right) \\
&= \tilde{w}h_{\hat{\lambda}}\left(\hat{\lambda}_{old}, \phi_{old}\right) \\
&> \tilde{w}h_{\hat{\lambda}}\left(\hat{\lambda}^*, \phi_{old}\right) \\
&= \tilde{w}h_{\hat{\lambda}}\left(\hat{\lambda}^*, \phi^*\right)
\end{aligned}
$$

which is again a contradiction.

Case 3. Finally assume that $\hat{\lambda}\left(\hat{q}\right)$ and $\phi\left(\hat{q}\right)$ both are increasing functions. Concavity and boundedness of v implies $\lim_{\hat{q}\to\infty} v'\left(\hat{q}\right) = 0$ and accordingly $\lim_{\hat{q}\to\infty} \frac{\beta\phi(\hat{q}^{**})}{1+g}v'\left(\frac{\hat{q}^{**}+\hat{\lambda}(\hat{q}^{**})}{1+g}\right) = 0$. Take $\hat{q}^* \in Q$ and let $\xi \equiv \tilde{w}h_{\hat{\lambda}}\left(\hat{\lambda}\left(\hat{q}^*\right), \phi\left(\hat{q}^*\right)\right)$. Then $\exists \hat{q}^{**} > \hat{q}^* \in Q$ such that $\frac{\beta\phi(\hat{q}^{**})}{1+g}v'\left(\frac{\hat{q}^{**}+\hat{\lambda}(\hat{q}^{**})}{1+g}\right) < \xi \leq \tilde{w}h_{\hat{\lambda}}\left(\hat{\lambda}\left(\hat{q}^{**}\right), \phi\left(\hat{q}^{**}\right)\right)$ where the first inequality uses the limit condition and the second inequality uses the monotonicity of the R&D decisions. The result is a contradiction to the fact that $\left(\hat{\lambda}\left(\hat{q}^{**}\right), \phi\left(\hat{q}^{**}\right)\right)$ is an equilibrium. Then it must be the case that $\hat{\lambda}\left(\hat{q}\right)$ or $\phi\left(\hat{q}\right)$ is strictly decreasing for some $\hat{q}^* < \hat{q}^{**}$. Since it was proven under Case1-2 that $\hat{\lambda}\left(\hat{q}\right)$ and $\phi\left(\hat{q}\right)$ cannot move in the opposite directions ever, the only possibility is that both $\hat{\lambda}\left(\hat{q}\right)$ are $\phi\left(\hat{q}\right)$ strictly decreasing functions.

QED.

∎

Proof of Theorem (3.2). *i*) Theorem 3.1 shows that $\hat{\lambda}\left(\hat{q}\right)$ is decreasing such that $\hat{\lambda}\left(\hat{q}\right) \to 0$ as $\hat{q} \to \infty$. Since $g^*, \hat{q}_{max} > 0$, the result follows.

ii) Next period's technology satisfies the following inequality, $\hat{q}_{t+t} < \frac{q_t + \lambda}{w(1+g)}$ since the proba-

bility of failure is not taken into account. Rearranging the same expression gives $\hat{q}_{t+t} + g^*\hat{q}_{t+t} <$ $\hat{q}_t + \hat{\lambda}$. Utilizing part (i) delivers the desired result. ∎

Derivation of (3.24) . Recall the definition of the technology index, $Q_t \equiv \int_I q_{i,t} di$. Its growth rate can be expressed as

$$Q_{t+1} = \rho \left(q_{\min,t+1} + \bar{q}_{t+1}\right)/2 + (1 - \rho) \int_I \left[\phi_{i,t}\left(q_{i,t} + \lambda_{i,t}\right) + \left(1 - \phi_{i,t}\right) q_{i,t}\right] di$$

Then

$$
\begin{aligned}
g &= Q_{t+1}/Q_t - 1 \\
&= \frac{(1 + g)\rho\left(q_{\min,t} + \bar{q}_t\right)/2 + (1 - \rho)\int_I \left[\phi_{i,t}\left(q_{i,t} + \lambda_{i,t}\right) + \left(1 - \phi_{i,t}\right) q_{i,t}\right] di}{\int_I q_{i,t} di} - 1 \\
&= \frac{(1 - \rho)\int_I \left[\phi_{i,t}\left(q_{i,t} + \lambda_{i,t}\right) + \left(1 - \phi_{i,t}\right) q_{i,t}\right] di}{\int_I q_i di - \rho\left(q_{\min,t} + \bar{q}_t\right)/2} - 1
\end{aligned}
$$

In this expression, the second line uses the fact that the minimum and the average technologies grow at the steady-state growth rate g. Some simple algebra leads to the third line. ∎

Proof of Proposition (3.2). $\phi\left(\hat{q}\right) \in [0, 1]$ is clearly bounded. Moreover, $0 < \phi_{\min} \leq \phi\left(\hat{S}\right)$ thanks to the Inada conditions and the definition of upper hemi-continuity (Mas-Colell, Whinston and Green, 1995, p.950). Let's consider $\hat{q}_{out}$ as the state where all exiters are collected. Assume $g^* > 0$. Then $\exists n^* \equiv \min\{n \in \mathbb{N}\} \geq 1$ such that

$$\frac{\hat{q}_{\max}}{(1 + g)^n} < \hat{q}_{\min}$$

equivalently

$$\frac{\log\left(\hat{q}_{\max}\right) - \log\left(\hat{q}_{\min}\right)}{\log\left(1 + g\right)} < n.$$

Then for any $\hat{q} \in \hat{Q}$

$$P^{n^*}\left(\hat{q}, \{\hat{q}_{out}\}\right) \geq \left(1 - \phi_{\max}\right)^{n^*} > \epsilon > 0$$

152

Since for all $A \subset 2^{\hat{Q}}$, either $\hat{q}_{out} \in A$ or $\hat{q}_{out} \notin A$, we have that

$$P^{n^*}(\hat{q}, A) \geq P^{n^*}(\hat{q}, \{\hat{q}_{out}\}) > \epsilon$$

or

$$P^{n^*}(\hat{q}, A^c) \geq P^{n^*}(\hat{q}, \{\hat{q}_{out}\}) > \epsilon.$$

This proves the existence of $\epsilon > 0$, an integer $n^* \geq 1$ such that for any $A \subset 2^{\hat{Q}}$, either $P^{n^*}(\hat{q}, A) \geq \epsilon$ or $P^{n^*}(\hat{q}, A^c) \geq \epsilon \ \forall \hat{q} \in \hat{Q}$. This is *Condition M* in Stokey-Lucas (1989) p.348.

Then the existence of the invariant distribution follows from Theorem 11.12 in Stokey-Lucas (1989 p.350).

If $g^* = 0$, then the invariant distribution consists of two atoms at $\hat{q}_{out}$ and $\hat{q}_{min}$. ∎

Proof of Proposition (3.3). Recall that $q_i \geq q_{min} \in \mathbb{R}_{++}$. In this economy, the equilibrium consists of the firm's maximization problem where it takes the following values as given:

$$m \equiv [K, \tilde{w}, g]$$

where $K \equiv \bar{c} + \gamma \bar{y}/Y$. Therefore to prove that an equilibrium exists, we need to show that the equilibrium values of these variables exist. Let $m \in \mathbf{M}$ be a generic argument of the following operator

$$\Phi(m) : \mathbf{M} \to \mathbf{M} \tag{3.32}$$

such that

$$\Phi(m) \equiv [\Phi_K(m), \Phi_{\tilde{w}}(m), \Phi_g(m)]^T$$

and $\mathbf{M} \equiv [0, K_{max}] \times [0, 1] \times [0, g_{max}]$. The equilibrium will be shown as a fixed point of this operator.

Let $Y_{\max}$ be an upper bound for the values that the final good can take. I will define $Y_{\max}$ as the social planner's solution to the static problem. The planner would use infinitely many firms to exploit the concavity of the production function. Therefore the second term in (3.7) would drop. As a result, the static maximization problem of the planner becomes

$$\max_{D_{sp}} \left\{ A D_{sp} - \frac{\delta}{2} D_{sp}^2 \right\}$$

where we also make use of the fact that all the labor is employed *only* for production to maximize the final output. The solution implies

$$D_{sp} = \frac{A}{\delta} \text{ and } Y_{\max} \equiv Y_{sp} = \frac{A^2}{2\delta}.$$

Accordingly, we can define $c_i = \frac{w}{q_i} < \frac{w}{q_{\min}} < \frac{Y_{\max}}{q_{\min}} \equiv c_{\max}$. Now we are ready to define the operator $\boldsymbol{\Phi}$.

Consider the first argument, K. It is simply a weigted summation of the average marginal cost and output. Since both of these values are bounded from below by 0, the following (sub)operator is also bounded from below,

$$
\begin{aligned}
K &= \min \left\{ \bar{c} + \frac{\gamma \bar{y}}{Y}, \ K_{\max} \right\} \\
&\equiv \Phi_K(m)
\end{aligned}
$$

The upper bound comes from the fact that $\pi_{\max} < Y$ which implies $\tilde{\pi}_{\max} < 1$. Therefore $1 = \tilde{\pi}_{\max} > \tilde{\pi}_i = \frac{1}{2\gamma}(K - c_i)^2$. Then

$$\sqrt{2\gamma} + c_{\max} \equiv K_{\max} > K.$$

Similarly $K > 0$ as long as $\exists i \in I$ such that $c_i < \bar{c} + \frac{\gamma \bar{y}}{Y}$. This is violated iff $c_i = \bar{c}$, $\forall i \in I$ and

154

$N = \infty$. However, the free entry condition rules this out. Therefore $0 < K < K_{\max}$.

Next, consider the (sub)operator for the labor share defined as

$$\tilde{w} = \min\left\{ N\left(\frac{\bar{y}}{Y}\overline{\left(\frac{1}{\hat{q}}\right)} + \frac{1}{\gamma}\left[\overline{\left(\frac{1}{\hat{q}}\right)}^2 - \overline{\left(\frac{1}{\hat{q}^2}\right)}\right]\right) + \tilde{w}N\bar{h},\ 1\right\}$$
$$\equiv \Phi_{\tilde{w}}(m)$$

which simply utilizes the labor market clearing condition. We need to show that $\tilde{w} < 1$ to ensure labor market clearing. Now assume $\tilde{w} = 1$. This means that all the output is being paid to labor as wage leaving the firms with 0 profits. However, in that case there will be no demand for labor since the firms will choose to exit and no firm will enter the market due to non-zero outside option and entry fee. This contradicting the fact that $\tilde{w} = 1$ would clear the market. Therefore $\tilde{w} < 1$ will clear the market.

The third operator for the growth rate is defined from equation (3.24)

$$g = \min\left\{ \frac{(1-\rho)\int_{I^*}\left(\hat{\lambda}_i^*\phi_i^* + \hat{q}_i\right)di}{\int_{I^*}\hat{q}_i di - \rho\left(\hat{q}_{\min} + \bar{\hat{q}}\right)/2} - 1, g_{\max}\right\}$$
$$\equiv \Phi_g(m)$$

where $g_{\max} \equiv \frac{1}{q_{\min}} = c_{\max}$. Clearly the growth rate will never hit the upper limit due to the Inada conditions at $\phi, \hat{\lambda} = 1$. Therefore $g < g_{\max}$.

Note that the operator $\Phi(m)$ is an upper-hemicontinuous correspondence since the R&D decisions are upper-hemicontinuous from theorem (3.1). The set $\mathbf{M}$ is compact-convex valued , therefore the existence of the fixed point follows from Kakutani's Theorem in Mas-Colell, Whinston and Green (1995, $p.953$). ∎

Proof of Theorem (3.2).

R&D spending can be expressed as

$$R\&D\ Intensity \;=\; \frac{wh_i}{p_i y_i}$$

$$=\; \frac{wh_i}{\frac{1}{2\gamma}\left(K^2 - \left(\frac{1}{\hat{q}_{i,t}}\right)^2\right)}$$

Proposition 3.1 showed that both $\hat{\lambda}_i$ and ϕ_i are decreasing for $\hat{q}_i > \frac{3\hat{q}(1+g^*)}{2}$. On the other hand, the denominator is decreasing for $\hat{q}_i$. This completes the proof. ∎

Proof of Proposition (3.4). Recall that

$$y_i \;=\; \bar{y} + \frac{Y}{\gamma}\left(\bar{c} - \frac{w}{q_i}\right)$$

$$=\; \frac{Y}{\gamma}\left[K - \frac{w}{q_i}\right]$$

where $K \equiv \gamma\bar{y}/Y + \bar{c}$. The expected firm growth is

$$\mathfrak{G}\left(\hat{q}_i\right) \;=\; \eta\left(-1\right) + \left(1-\eta\right)\left[\frac{\left(1+g^*\right)Y}{\gamma}\frac{\left[K - \left(1+g^*\right)w\left[\frac{\phi}{q_i+\lambda_i} + \frac{1-\phi}{q_i}\right]\right]}{\frac{Y}{\gamma}\left[K^* - \frac{w}{q_i}\right]} - 1\right]$$

$$=\; -\eta + \left(1-\eta\right)\left[g^* + \left(1+g^*\right)\frac{1 + \left(\frac{\phi_i\lambda_i}{q_i+\lambda_i} - 1\right)\left(1+g^*\right)}{K^*\hat{q}_i - 1}\right] \qquad (3.33)$$

Note that the first line replaced the growth rate of exiters by -100%. ∎

Proof of Theorem (3.3).

Since the denominator $\left(K^*\hat{q}_i - 1\right)$ is increasing in $\hat{q}_i$, it is sufficient to show that $\frac{\phi_i\hat{\lambda}_i}{\hat{q}_i+\hat{\lambda}_i}$ is decreasing in $\hat{q}_i$. Taking the derivative wrt $\hat{q}_i$ we get

$$\frac{d}{d\hat{q}_i}\left(\frac{\phi_i\hat{\lambda}_i}{\hat{q}_i+\hat{\lambda}_i}\right) \;=\; \frac{\left[\hat{\lambda}'\left(\hat{q}_i\right)\phi\left(\hat{q}_i\right) + \hat{\lambda}\left(\hat{q}_i\right)\phi'\left(\hat{q}_i\right)\right]\left[\hat{q}_i + \hat{\lambda}\left(\hat{q}_i\right)\right] - \left[1 + \hat{\lambda}'\left(\hat{q}_i\right)\right]\hat{\lambda}\left(\hat{q}_i\right)\phi\left(\hat{q}_i\right)}{\left[\hat{q}_i + \hat{\lambda}\left(\hat{q}_i\right)\right]^2}$$

$$=\; \frac{\hat{\lambda}'\left(\hat{q}_i\right)\phi\left(\hat{q}_i\right)\hat{q}_i + \hat{\lambda}\left(\hat{q}_i\right)\phi'\left(\hat{q}_i\right)\hat{q}_i + \hat{\lambda}^2\left(\hat{q}_i\right)\phi'\left(\hat{q}_i\right) - \hat{\lambda}\left(\hat{q}_i\right)\phi\left(\hat{q}_i\right)}{\left[\hat{q}_i + \hat{\lambda}\left(\hat{q}_i\right)\right]^2}$$

$$<\; 0$$

156

Therefore the expected growth rate is deacresing in $\hat{q}_i$, equivalently in firm size. ∎

3.9 References

Acemoglu, Daron (2008) *Introduction to Modern Economic Growth,* Princeton University Press, Princeton, NJ.

Acemoglu, Daron and Akcigit, Ufuk (2008) "State Dependent Intellectual Property Rights Policy" *MIT mimeo.*

Adda, Jerome and Russell Cooper (2003) *Dynamic Economics: Quantitative Methods and Applications,* MIT Press, Cambridge MA.

Aghion, Philippe, Christopher Harris and John Vickers (1997) "Competition and Growth with Step-by-Step Innovation: An Example" *European Economic Review,* 41, 771-782.

Aghion, Philippe, Christopher Harris, Peter Howitt and John Vickers (2001) "Competition, Imitation and Growth with Step-by-Step Innovation" *Review of Economic Studies,* 68(3), 467-492.

Anselin, Luc, Atilla Varga, and Zoltan Acs (1997) "Local Geographic Spillovers between University Research and High Technology Innovations" *Journal of Urban Economics* 42(3), 422-448

Anselin, Luc, Atilla Varga, and Zoltan Acs (2000) "Geographical spillovers and university research: A spatial econometric perspective." *Growth and Change* 31, 501-15.

Bloom, Nicholas, Rachel Griffith and Alexander Klemm (2001) "Issues in the Design and Implementation of an R&D Tax Credit for UK Firms" *The Institute of Fiscal Studies Briefing Notes,* No. 15.

Bloom, Nicholas, (2009) "The Impact of Uncertainty Shocks", *Econometrica,* forthcoming

Campbell, Jeffrey R. and Hugo A. Hopenhayn (2002) "Market Size Matters" *NBER Working Papers*, No 9113.

Corcos, Gregory, Massimo Del Gatto, Giordano Mion and Gianmarco I.P. Ottaviano (2007) "Productivity and Firm Selection: Intra- vs International Trade" *CORE Discussion Paper* No. 2007/60.

Goffe, William (1996) "SIMANN: A Global Optimization Algorithm using Simulated Annealing" *Studies in Nonlinear Dynamics & Econometrics*, Berkeley Electronic Press, 1(3), 169-176.

Gourieroux, Chriastian and Alain Monfort (1996) "Simulation-Based Econometric Methods" Oxford University Press.

Griliches, Zvi (1992) "The Search for R&D Spillovers" *Scandinavian Journal of Economics*, 94, 29–47.

Griliches, Zvi (1995) "R&D and Productivity: Econometric Results and Measurement Issues" in P. Stoneman, ed. *Handbook of the Economics of Innovation and Technological Change*, Oxford, Blackwell, 52-89.

Guner, Nezih, Gustavo Ventura and Yi Xu (2008) "Macroeconomic Implications of Size-dependent Policies" *Review of Economic Dynamics*, 11(4), 721-744.

Hall, Robert E. and Charles Jones (1999) "Why Do Some Countries Produce So Much More Output per Worker Than Others?" *Quarterly Journal of Economics*, 114 (1), 83–116.

Jaffe, Adam B., Manuel Trajtenberg and Rebecca Henderson (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations" *The Quarterly Journal of Economics*, 108(3), 577-98.

Jones, Charles I. and John C. Williams (1998) "Measuring the Social Return to

R&D" *Quarterly Journal of Economics*, 113(4), 1119-1135.

Klenow, Pete and Andres Rodriguez-Clare (1997) "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" in B. Bernanke and J. Rotemberg, Editors, *NBER Macro Annual*, MIT Press.

Mas-Colell, Andreu, Michael Whinston and Jerry Green (1995) *Microeconomic Theory*, Oxford University Press.

Melitz, Marc J. and Gianmarco I.P. Ottaviano (2005) "Market Size, Trade, and Productivity" *Review of Economic Studies*, 75(1), 295-316.

Ottaviano, Gianmarco, Takatoshi Tabuchi and Jacques-François Thisse (2002) "Agglomeration and Trade Revisited" *International Economic Review*, 43, 409-436.

Restuccia, Diego and Richard Rogerson (2007) "Policy Distortions and Aggregate Productivity with Heterogeneous Plants" *NBER Working Paper*, No. 13018.

Segerstrom, Paul (1998) "Endogenous Growth without Scale Effects" *American Economic Review*, 88(5) 1290-1310 .

Spence, Michael (1984) "Cost Reduction, Competition, and Industry Performance" *Econometrica*, 52(1), 101-21.

Stokey, Nancy. L., Robert. E. Lucas and Edward C. Prescott (1989) "Recursive Methods in Economic Dynamics" Harvard University Press.

Syverson, Chad (2004) "Market Structure and Productivity: A Concrete Example" *Journal of Political Economy*, 112 (6), 1181-1222.

Syverson, Chad (2007) "Prices, Spatial Competition, and Heterogeneous Producers: An Empirical Test" *Journal of Industrial Economics*, 55(2), 197-222.