Knowledge Spillovers and Firm Size Heterogeneity[1]

Jie (April) Cai[2]
University of British Columbia

August 4, 2009

Abstract: This paper explains two data facts related to firm size distribution. First, it uses sector-specific inter-firm knowledge spillovers to explain the sectoral differences in firm size heterogeneity. I formalize an environment in which greater inter-firm knowledge spillovers in a sector induce firms in that sector to invest relatively more in imitation. This implies that imitation contributes a greater share to firm growth rates in these sectors. Greater imitation also causes faster catch-up by lagging firms and declining firm growth rate with firm size. Hence, the sectoral firm size distribution becomes more homogeneous in the sectors with greater knowledge spillovers. Second, in a multi-sector version of this environment, I use inter-sector knowledge spillovers to explain the observed dependent *Pareto*[3] size distributions in every subset of the economy. I test the model using patent citation data and find support for both its sectoral and aggregate predictions.

JEL Classification: O33 O41 L11 L6

Keywords: inter-firm knowledge spillovers, firm size heterogeneity, sectoral differences, scale-dependent firm growth rate, inter-sector knowledge spillover, *Pareto* distribution

1 Introduction

More and more firm- or establishment-level data show that firm size distributions within narrowly defined sectors and within the overall economy are widely dispersed and follow a *Pareto* distribution. Two important related questions are not well understood in the literature of firm growth dynamics. First, why is firm size heterogeneity[4] different across sectors? And second, why does *Pareto* firm size distribution exists in every subset of the economy? Moreover, why are firm size variables in different sectors dependent on each other?

This paper uses intra-sector and inter-sector knowledge spillovers, respectively, to answer these two questions. In a one-sector model, cross-sector difference in firm size heterogeneity can be attributed to sector-specific intra-sector knowledge diffusion efficiency. The one-sector model extends the endogenous innovation model of Klette and Kortum (2004) by giving firms the option to imitate. In sectors with more abundant knowledge spillovers, firms invest relatively more in imitation, as compared to innovation. Imitation then contributes a greater share to gross growth rate. Since the equal opportunity to learn provides a stronger impetus to small firms, firm growth rate drops faster as the firm becomes larger. The sectoral firm size distribution is more homogeneous if

[1] I would like to thank my supervisor, Amartya Lahiri, and committee members, Paul Beaudry and Patrick Francois, for their advice, guidance and continuous encouragement. I would also like to thank Matilde Bombardini, Michael Devereux, Wenhui Fan, Keith Head, Ran Jing, Henry Siu and seminar attendants at Auckland University, Brock University, Ohio State University and University of New South Wales for their contributive suggestions. Thanks Luca David Opromolla for the Chilean Manufacturing Data.

[2] University of British Columbia, Department of Economics, 997 - 1873 East Mall, Vancouver BC, Canada, V6T 1Z1. aprilcai@interchange.ubc.ca. http://grad.econ.ubc.ca/aprilcai

[3] If X is a random variable with a Pareto distribution, then the probability that X is greater than some number x is given by $Pr(X > x) = \left(\frac{x}{x_m}\right)^{-\mu}$, for all $x \geq x_m$, where x_m is the (necessarily positive) minimum possible value of X, and is a positive parameter. Pareto distribution is the continuous version of power law distribution. Zipf's law is a power law distribution with exponent $\mu = 1$, at least approximately.

[4] Using French, Chilean and U.S. firm-level data, Appendix A shows that sector-specific firm size heterogeneity is robust to different proxies of firm size, stable over time in the same country, and highly correlated across different countries in the same year.

small firms have more opportunities to catch up with the leaders. The model implications are confirmed using NBER Patent Citation Data, which provides a measure of knowledge spillovers and the appropriate information for distinguishing the contributions by imitation and innovation to firm growth rate. The one-sector model also provides micro-foundation for an optimal R&D policy tailored for every sector.

A multi-sector model incorporates two additional facts absent in the one-sector model: firms develop products in multiple sectors and inter-sector knowledge spillovers integrates the firm growth dynamics in all sectors. Firm growth dynamics in any subset of the economy are subject to a similar influence from all sectors, and the firm size distribution therefore converges universally to a *Pareto* distribution, not only within narrowly defined sectors but in the economy overall. The open economy version of the multi-sector model suggests favorable trade policy for the sectors that have abundant intra-sector knowledge spillovers and contribute inter-sector knowledge spillovers to other sectors.

The one-sector model in this paper extends that of Klette and Kortum (2004) by allowing firms to create new goods by imitation, in addition to innovation. The difference between innovation and imitation is that innovation relies on a firm's private knowledge (measured by its current number of goods), while imitation relies on a sector's public knowledge (measured by the average firm size in the sector). Both types of R&D are subject to identically independent (i. i. d.) shocks, which are necessary to induce the *Pareto* firm size distribution. The sector-specific knowledge spillovers efficiency is given by a firm's productivity in utilizing private knowledge in innovation and public knowledge in imitation. When it is relatively more efficient to imitate than innovate, firms invest relatively more in imitation; as a result, the imitation rate contributes a greater share to the overall growth rate for the firm and the entire sector.

The scale-dependent firm's growth rate is the summation of the scale-independent innovation rate and the scale-dependent imitation rate. The innovation rate is independent of firm size, because, as specified in the Cobb-Douglass production function of new goods, output (innovated new goods) is proportional to the input (firm's private knowledge capital). In contrast, the imitation rate decreases with firm size, because the output (imitated new goods) is proportional to the input (public knowledge pool). Smaller firms have a higher growth rate caused by imitation, when the output is divided by firm size. In sectors with higher imitation efficiency, firms' growth rates drop faster as the firms become bigger, since the imitation accounts for a greater share of total growth rate. A larger growth rate gap between smaller and larger firms causes faster firm size mean reversion and a more homogeneous firm size distribution.

According to Kesten (1973), when the innovation shock and imitation shock are i. i. d. across time and firm, firm size distribution within the sector converges to a *Pareto* distribution with scale parameter μ. While in Klette and Kortum's (2004) environment without imitation, the firm size distribution converges to a logarithmic distribution. When the innovation risk follows log-normal distribution, the closed form solution of the firm size heterogeneity measure, $1/\mu$, increases with the volatility of innovation shock and decreases with imitation's contribution to gross growth rate of the sector. Intuitively, the innovation shocks underlay the origin of firm size dispersion, while imitation shocks alleviate firms' size differences by allowing firms to learn and catch up.

The firm growth dynamics model with both innovation shock and imitation shock help understand a recent discovery about the declining small private firms' volatility and rising large public firms's volatility (volatility convergence) in Comin and Mulani (2006, 2007) and Davis, Haltiwanger, Jarmin and Miranda (2006). The model shows that innovation risk is the major component of large firm's volatility, while imitation risk is the driving force of small firm's volatility. When some policy or technological changes encouraged firms to invest more in innovation and less in imitation, more (less) risky projects are taken in innovation (imitation). As a result, a large firm's volatility increases with its innovation risk, while a small firm's volatility declines with its imitation risk. To my knowledge, this paper is the first theoretical attempt to explain the above fact.

The one-sector model has three testable implications. First, scale-dependent firm growth rate arises purely from scale-dependent imitation rate. Specifically, a surviving firm's imitation rate drops as its size increases, while a firm's innovation rate is independent of firm size. Moreover, a firm's growth rate is more scale-dependent in sectors with more abundant knowledge spillovers and more homogeneous firm size. Second, firm size heterogeneity decreases with imitation's contribution to gross growth rate, and increases with the volatility of innovation risk. Third, knowledge diffuses faster in sectors with a more homogeneous firm size distribution

The challenge in testing the first two implications is to distinguish imitation rate from innovation rate in the total firm growth rate. This is done by differentiating between citations given to the citing firm's old patents (inside citations) and citations given to other firms' old patents (outside citations). The total firm growth rate is split into innovation rate and imitation rate according to the ratio between inside citations and outside citations.

In the regression of innovation (imitation) rate on firm size, the coefficient is barely (always) statistically significantly different from zero among 42 sectors, as predicted by the second model implication. This result supports the idea that imitation rate declines with firm size, while innovation rate is independent of firm size. Moreover, the regression coefficient of imitation rate on firm size decreases with firm size heterogeneity, which confirms the model prediction that imitation rate is more scale-dependent in sectors with more homogeneous firm size distribution.

With estimated innovation rate and imitation rate for every firm within a sector, I can derive the sector-level imitation rate's share in gross growth rate and the variance of the log scale innovation rate. The second implication is also supported by the data: firm size heterogeneity is negatively related to imitation rate's share in gross growth rate and positively related to the volatility of innovation risk.

To test the third implication, I employ within-sector U.S. patent citations in NBER Patent Citation Data as a measure of intra-sector knowledge spillovers. The cross-firm knowledge diffusion speed and the percentage of cross-firm citations among all citations are negatively correlated with the firm size distribution heterogeneity. In the regressions, I control the geographic distance between the citing and the cited patent, size of the citing organization, cited organization and sector size. The knowledge spillovers speed is measured by citation time lag[5]. As a robustness check, the cross-sector difference in knowledge spillovers also holds in the citation data including all G7 countries.

Now, to turn to the second question: why does firm size distribution follow a *Pareto* distribution not only within narrowly defined sectors but also in the entire economy? According to Jessen and Mikosch (2006), summation or pooling of independent *Pareto*-distributed variables induces a new *Pareto*-distributed variable. However, the scale parameter of the new *Pareto* distribution should be equal to the smallest scale parameter of the component distributions. In contrast, firm-level data (Figures 1 and 2) show that the scale parameter of the size distribution for the whole economy is in the middle of the range of the component distributions' scale parameters. Therefore, some mechanism must make firm size and firm growth dynamics in different sectors dependent.

The multi-sector model adds two important elements to the one-sector model. First, many firms develop products in multiple sectors. In the NBER Patent database, on average every organization applied for patents in 6.5 out of 42 patent categories; moreover, larger organizations cover more categories (see Table 3). Second, inter-sector knowledge spillovers is as important as within-sector knowledge diffusion. Inter-sector citations amount to 37% of total citations in the citation data. In the multi-sector model, when firms invent new products in a single sector, they can apply their private knowledge capital from all sectors. Also, a firm's growth in a single sector is affected by its previous knowledge capital in all sectors. As a result, a firm's overall size, which is the

[5] The citation lag, the time difference between the application time of the citing patent and that of the cited patent, indicates the time needed for the knowledge to travel between the citing patent inventor and the cited patent inventor.

summation of its branches in all sectors, is influenced by its private knowledge capital in all sectors. Since a firm's growth dynamics in any subset of the economy follows a similar formula, firm size distribution converges to *Pareto* distribution universally in any subset of the economy.

Thanks to the endogenous growth framework, these micro-founded firm-level models also provide insight into national and international economic growth problems. In a close economy, with two types of R&D, the optimal growth policy through R&D support is sector-specific. When knowledge diffusion is faster within the borders of a firm than across, subsidizing innovation by 1% induces a higher growth increase compared to fostering imitation by the same amount, and vice versa. The reason is that firms allocate more inputs to the R&D type that is more efficient in knowledge diffusion, and therefore R&D output exhibits increasing returns to knowledge diffusion productivity. Depending on the sector, it is thus optimal to implement a policy favorable to the type of R&D that has a comparative advantage in knowledge diffusion.

In future research, I show in an open economy version of this paper that trade liberalization causes a higher sectoral growth rate in sectors with more abundant intra-sector knowledge spillovers. Moreover, with multiple sectors, the entire economy can benefit from lowering the trade barriers of the sectors that generates inter-sector knowledge spillovers. One implication is that countries specializing in different sectors may exhibit different growth performance in globalization.

1.1 Literature Review

In the firm growth dynamics literature, I have identified two strands in the theoretical debate about sources of firm size heterogeneity. The first strand, represented by the work of Lucas (1978), Jovanovic (1992), and Klette and Raknerud (2002), emphasizes the role of a manager's various talents in creating permanent differences in firm efficiency. The second strand, elaborated by Hopenhayn (1992), Ericson and Pakes (1995), Klepper (1996), Klette and Kortum (2004), Klepper and Thompson (2007), and Luttmer (2008), contends that firm size dispersion is caused by accumulated idiosyncratic shocks over a firm's life cycle. Seker (2008) incorporates both of these strands and tries to distinguish the contribution of each.

In addition to exploring the origins of firm size heterogeneity, the literature on firm size dynamics tries to explain interesting stylized facts observed in firm-level data. First, the firm size distribution follows a *Pareto* distribution, both within one-sector and in the entire economy. Meanwhile, the heterogeneity of firm size distribution varies across sectors (see Figures 1 and 2), as documented by Axtell (2001); Helpman, Melitz and Yeaple (2004); Rossi-Hansberg and Wright (2007) ("RW" henceforth); and Luttmer (2008). Second, a surviving firm's expected growth rate drops with firm size, or put differently, firm growth rate is scale-dependent.[6]

Studies take various approaches when modeling scale-dependent surviving firm growth rate. Cooley and Quadini (2001), Cabral and Mata (2003), Albuquerque and Hopenhayn (2004) and Clementi and Hopenhayn (2006) show that financial market friction can induce firm growth rate to decline with firm size. In Klette and Kortum (2004), every firm has the same unconditional growth rate, while small firms have a higher growth rate conditional on survival, because they are less likely to survive than big firms. Selection is the key to having scale-dependent firm growth rate. Klepper and Thompson (2007) use creation and deconstruction of submarkets to explain firm size dynamics. The size decrement due to submarket deconstruction is proportional to firm size, while the size increment due to the creation of emerging markets is independent of firm size. As the firm operates in more submarkets, the expected proportional increase in size declines. RW shows that small firms grow faster because households want to accumulate industry-specific human capital more rapidly in small firms where the

[6] See Evans (1987); Hall (1987); Dunne, Roberts and Samuelson (1989); Sutton (1997); Klette and Kortum (2004); and Luttmer(2008). Again, there are cross-sector differences in growth rate scale dependency, as demonstrated in RW.

marginal return of human capital is still high. Luttmer (2008) assumes that new firms enter with a high-quality blueprint, but that the blueprint's quality depreciates and becomes obsolete over time. Hence, firms choose to replicate their blueprints faster when they are smaller and their blueprints' quality is still high.

To my knowledge, only RW explain cross-sector differences in firm growth rate scale dependence and firm size heterogeneity in terms of sector-specific capital intensiveness. The firm growth rate drops faster in more capital intensive or less human capital intensive sectors because the marginal return to human capital decreases faster there. When small firms are more likely to catch up with big firms, firm size distribution becomes more homogeneous. Capital intensiveness can explain cross-sector differences for broad sector divisions (for instance, between education and construction and manufacturing, etc.), but it fails to explain differences across a more refined division of manufacturing sectors. The patent citation data used in this paper primarily covers manufacturing sectors. It shows that examining knowledge spillovers efficiency is a more promising way to account for differences across manufacturing sectors.

Luttmer (2007) and RW also consider the role of knowledge spillovers in shaping firm size distribution. In Luttmer (2007), only entrants learn from incumbents, as new entrants can learn more from incumbent, the firm size distribution is more homogeneous. In the learning-by-doing (LBD) extension of RW, the industrial total output enters the accumulation function of industry-specific human capital. The authors show that a larger externality in LBD leads to a faster mean reversion in human capital stock and a more homogeneous firm size distribution. The conclusion of the extension is that capital intensiveness and the LBD externality jointly determine firm size dispersion. In addition, the firm size distribution in their paper converges to a log-normal distribution, instead of a *Pareto* distribution. Unfortunately, unlike this paper, neither of the above two papers provides empirical evidence to support its theoretical predictions on knowledge spillovers.

This paper differs from Luttmer (2007) in that it allows every firm, instead of just new entrants, to imitate. It is closer to RW's extension featuring LBD externality, but this paper uses a micro-founded approach, while RW use a macroeconomic approach. In some sense, knowledge spillovers is one reason for the depreciation of a blueprint's quality in Luttmer (2008). Expecting that others will "steal" its blueprint in the future, the owner of a blueprint chooses to replicate it faster before others imitate it.

To the best of my knowledge, no research has studied the universal *Pareto* firm size distribution that is found across all subset of the economy. In addition, this paper makes a distinct contribution by providing policy suggestions concerning the sector-specific optimal R&D policy and trade policy.

The remainder of the paper is organized as follows. In section 2, the one-sector model shows that sector-specific intra-sector knowledge diffusion efficiency determines a firm's choice of endogenous innovation and imitation inputs, which affects in turn the firm size heterogeneity. I test in section 3 the implications of the one-sector model with NBER Patent Citation Data. The multi-sector model is presented in section 4, where I show that inter-sector knowledge spillovers integrates growth dynamics in all sectors and induces a *Pareto* size distribution in all subsets of the economy. Section 5 concludes.

2 One-Sector Model

2.1 Consumer

The representative consumer faces the following problem:

$$U = \max_{\{x_{i,t}\}} \int_0^\infty \rho^t \left[\log\left(Y_t\right)\right] dt \tag{1}$$

$$Y_t = \left(\int_0^{I_t} x_{i,t}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}.$$

ρ is the time preference of the representative consumer; Y_t is the consumption of final goods; x_{it} is the consumption of intermediate good i; product i is sold at price $p_{i,t}$. P_t is the aggregate price index. There are I_t intermediate goods in the economy at time t. $\sigma > 1$ is the elasticity of substitution between intermediate goods. Consumer demand for intermediate goods is

$$x_{i,t} = Y_t \left(\frac{P_t}{p_{it}} \right)^{\sigma}$$

$$P_t = \left(\int_0^{I_t} p_{it}^{1-\sigma} di \right)^{\frac{1}{1-\sigma}}$$

2.2 Firms

There is only one sector with T firms in the economy. T is a large number so that each firm is tiny relative to the economy. Firm f hires one unit of production labor to produce one unit of goods. The wage rate is the numeraire. According to Dixit and Stiglitz (1977), the profit-maximizing price for every product is $\frac{\sigma}{\sigma-1}$. In the monopolistic competitive market, the profit from each product is $\frac{1}{\sigma}\frac{Y_t}{I_t}$. Firm f produces $z_{f,t}$ number of goods that it has invented by time t. The total number of goods in the economy is $I_t = \int_0^{M_F} z_{f,t} df$.

Firms grow by inventing new goods. Firm f invents new goods by two types of R&D: innovation and imitation. Innovation uses firm's private knowledge capital and $N_{f,t}$ units of research hour. Here the size of firm f's private knowledge capital is measured by $z_{f,t}$ to represents firm f's experience in R&D. In contrast, imitation uses public knowledge capital $\bar{Z}_t$[7] and $M_{f,t}$ units of research hour. The size of public knowledge pool is measured by the average firm size $\bar{Z}_t$ in the industry[8].

Imitation here does not refer to simple reverse engineering and replication, it means improvements and upgrades of other firms' products. This assumption reflects the fact that patent law does not acknowledge simple replication. In order to gain a new patent, a firm must upgrade existing patented goods to demonstrate enough originality and creativity. Firms' private knowledge diffuses to the public knowledge pool through many channels.[9] Firms may or may not voluntarily reveal their private knowledge to the public, but interactions between firms always generate a steady flow of knowledge from firms' private pools to the public pool.

I borrow the Cobb-Douglass production function from Klette and Kortum (2004) to describe the new goods production function. The expected number of new goods depends on the amount of hours invested in research and knowledge capital.

$$E\left(\Delta z_{f,t}^N\right) = A_N N_{f,t}^{\alpha} z_{f,t}^{1-\alpha}$$

[7] In Appendix D, firms uses both private and public knowledge to imitate.

[8] This assumption implies learning is time consuming and no firm can afford acquiring all outside knowledge in one period, given that every firm is tiny relative to the entire sector. In other words, the inter-firm learning happens one-to-one instead of one-to-all. A pair of Firms are randomly matched to learn from each other. What one firm expects to learn from a random peer is the average firm size in the sector. In Appendix D, the matching is positive assortative: a larger firm on average learns from a larger peer.

This assumption also ensures that in the general equilibrium, economic growth rate is a constant independent of population size.

[9] In Duguet and MacGarvie (2005), there are 12 channels listed: external R&D, cooperative R&D, patents and licenses, analysis of competition, experts, equipment acquisition, hiring employees, communication with suppliers, communication with customers, mergers and acquisitions, joint ventures and alliances, and personnel exchange.

$$E\left(\Delta z_{f,t}^{M}\right) = A_M M_{f,t}^{\beta} \bar{Z}_t^{1-\beta}$$

$\Delta z_{f,t}^{N}$ ($\Delta z_{f,t}^{M}$) is the number of new goods invented by innovation (imitation). The productivity of innovation (imitation) A_N (A_M) is the sector-specific. This assumption captures the fact that technology is more standardized or codifiable in some sectors than others. Standardized industrial technology is easier to transplant across firms, while firm-specific technology is only suited for application within the inventing firm. Besides, standardized industrial technology facilitates workers to change employers within the same sector. If knowledge capital is embodied in workers, high labor turnover also helps to disseminate one firms' private knowledge to other firms.

Within a sector, the productivties in the two types of R&D A_N and A_M can be different, which implies that firms employ private and public knowledge at different costs. These costs include the searching cost of related existing knowledge; the reverse engineering cost of absorbing the existing knowledge; and the creation cost of adding novelty to the existing knowledge. Normally, firm borders block knowledge spillovers, and it is therefore more efficient to use private knowledge instead of public knowledge, i.e. $A_N > A_M$. [10]

The Cobb-Douglass knowledge production function hinges on two assumptions. First, R&D research hours $N_{f,t}$ and $M_{f,t}$ have decreasing marginal productivity. Second, with the same amount of innovative research hours, $N_{f,t}$, larger firms invent more new goods due to more R&D experience. Similarly, with the same amount of imitative research hours, $M_{f,t}$, firms with access to a deeper public knowledge pool $\bar{Z}_t$ create more new goods. When these two forces offset each other, the amount of research hours that each firm spends on innovation (imitation) is proportional to the size of private (public) knowledge pool $z_{f,t}$ ($\bar{Z}_t$).

Firm f chooses inputs in innovation $N_{f,t}$ and imitation $M_{f,t}$ to maximize its firm value $V(z_{f,t})$.

$$\max_{N_{f,t},\ M_{f,t}} \ V(z_{f,t}) = \frac{P_t Y_t}{\sigma} \frac{z_{f,t}}{I_t} + \rho E[V(z_{f,t+1})] - \frac{N_{f,t} + M_{f,t}}{\bar{Z}_t} \tag{2}$$

subject to

$$z_{f,t+1} = z_{f,t} + \Delta z_{f,t}^{N} + \Delta z_{f,t}^{M} \tag{3}$$

$$\frac{\Delta z_{f,t}^{N}}{z_{f,t}} = \frac{A_N N_{f,t}^{\alpha} z_{f,t}^{1-\alpha}}{z_{f,t}} + \varepsilon_{f,t}^{n} \tag{4}$$

$$\frac{\Delta z_{f,t}^{M}}{\bar{Z}_t} = \frac{A_M M_{f,t}^{\beta} \bar{Z}_t^{1-\beta}}{\bar{Z}_t} + \varepsilon_{f,t}^{m} \tag{5}$$

Firm f's investments in R&D decide the expected success rate of innovation and imitation, but the acutally realization of them in (4) and (5) are subject to i.i.d. shocks $\varepsilon_{f,t}^{n}$ and $\varepsilon_{f,t}^{m}$.[11] When firm f's manager chooses research inputs at the beginning of time t, she knows the distributions of $\varepsilon_{f,t}^{n}$ and $\varepsilon_{f,t}^{m}$ but not their actual realizations.

Firm f discounts future firm value by ρ. Labor productivity in R&D grows as fast as the public knowledge

[10] On average, firms tend to use private knowledge more frequently and sooner than public knowledge. In the NBER Patent Citation Data, on average every organization owns 0.17% of the old patent stock in the industry, but they cite over proportionally 11.1% of its own old patents. If the cited and citing patents are owned by the same organization, the average citation time lag is 5.86 years; otherwise the average time lag is 9.06 years. Citation lag is the application year of citing patent minus the application year of cited patent, which tells how long it takes the citing firm to acquire and make use of knowledge of the cited patent.

[11] $\varepsilon_{f,t}^{n}$ and $\varepsilon_{f,t}^{m}$ are zero mean random variables bounded from below, such that $\frac{\Delta z_{f,t}^{N}}{z_{f,t}}$ and $\frac{\Delta z_{f,t}^{M}}{\bar{Z}_t}$ are always positive.

pool size $\bar{Z}_t$. [12]Since each firm is tiny relative to the entire sector, firm f takes I_t, Y_t and P_t as given. I assume that firms get full liquidation value in case of exit, so that firm's current innovation and imitation decisions is independent of exit risk in the future.

One educated guess for the firm value is a linear function of the form: $V(z_{f,t}) = v\frac{z_{f,t}}{I_t} + u$. The first order conditions become:

$$N_{f,t} = \left(\frac{A_N \alpha v \rho I_t}{I_{t+1} M_F}\right)^{\frac{1}{1-\alpha}} z_{f,t} \tag{6}$$

$$M_{f,t} = \left(\frac{A_M \beta v \rho I_t}{I_{t+1} M_F}\right)^{\frac{1}{1-\beta}} \bar{Z}_t \tag{7}$$

$$v = \frac{P_t Y_t}{\sigma} + \frac{\rho v I_t}{I_{t+1}} \left[1 + (1-\alpha) A_N N_{f,t}^{\alpha} z_{f,t}^{-\alpha}\right] \tag{8}$$

(6) ((7)) equates the marginal cost of innovation (imitation) to the expected marginal return from innovation (imitation). A firm's optimal labor input in innovation $N_{f,t}$ is proportional to the firm's private knowledge $z_{f,t}$; and labor input in imitation $M_{f,t}$ is proportional to public knowledge $\bar{Z}_t$. (8) means that marginal value of current market share is the current marginal profit plus the discounted future profit from innovation. The constant component of the firm value function is given by

$$u = \frac{(1-\beta)\,\beta^{\frac{\beta}{1-\beta}} \left(\frac{A_M v \rho Y_t I_t}{Y_{t+1} I_{t+1} M_F}\right)^{\frac{1}{1-\beta}}}{1 - \frac{\rho Y_t}{Y_{t+1}}} \tag{9}$$

u is the expected discounted future profit from all the imitated products. In other words, u measures the public knowledge pool's externality to each firm. In the equilibrium with no entry, u must be smaller than or equal to the fixed entry cost F. Since potential entrants are firms with zero products, their expected profit from entering is purely the public knowledge externality u. When the externality u is greater than entry cost F, new entrants will keep entering and diluting the public knowledge pool $\bar{Z}_t$. Average firm size $\bar{Z}_t$ will shrink until u is equal to F and no more entry occurs.

Substituting (6) and (7) to (3), (4), and (5), the firm size dynamic process in (3) can be summarized by

$$z_{f,t+1} = R_{t+1} z_{f,t} + L_{f,t+1}, \tag{10}$$

where

$$R_{f,t+1} \equiv \frac{I_t}{I_{t+1}} \left(1 + A_N^{\frac{1}{1-\alpha}} \left[\frac{\alpha v \rho I_t}{I_{t+1} M_F}\right]^{\frac{\alpha}{1-\alpha}}\right) + \varepsilon_{f,t+1}^{n}, \tag{11}$$

$$L_{f,t+1} \equiv \frac{I_t}{I_{t+1}} A_M^{\frac{1}{1-\beta}} \left[\frac{\beta v \rho I_t}{I_{t+1} M_F}\right]^{\frac{\beta}{1-\beta}} + \varepsilon_{f,t+1}^{m}.$$

I decompose firm's expected growth rate $g_{f,t}$ into an innovation rate $r_{f,t}$ and an imitation rate $l_{f,t}$.

$$E\left(r_{f,t}\right) \equiv \frac{E\left(\Delta Z_{ft}^{N}\right)}{z_{f,t}} = A_N^{\frac{1}{1-\alpha}} \left[\frac{\alpha v \rho I_t}{I_{t+1} M_F}\right]^{\frac{\alpha}{1-\alpha}} \tag{12}$$

[12]This assumption keeps the number of R&D workers constant in general equilibrium while the number of goods can grow at a constant rate. Moreover, the endogenous growth rate of the economy is independent of the population size under this assumption.

$$E\left(l_{f,t}\right)\equiv\frac{E\left(\Delta Z_{ft}^{M}\right)}{z_{f,t}}=A_{M}^{\frac{1}{1-\beta}}\left[\frac{\beta\upsilon\rho I_{t}}{I_{t+1}M_{F}}\right]^{\frac{\beta}{1-\beta}}\frac{\bar{Z}_{t}}{z_{f,t}} \tag{13}$$

The expected innovation rate $E\left(r_{f,t}\right)$ is a constant and independent of firm size $z_{f,t}$; but the expected imitation rate $E\left(l_{f,t}\right)$ is scale-dependent. As firm f grows, its imitation rate declines simply because the public knowledge pool $\bar{Z}_t$ becomes smaller relative to firm f's size $z_{f,t}$. In total, the firm growth rate $g_{f,t}=r_{f,t}+l_{f,t}$ declines with firm size purely because of imitation rate.

Moreover, the expected imitation rate $E\left(r_{f,t}\right)$ declines faster (or is more scale-dependent), when cross-firm knowledge spillovers is more efficient (A_M is greater) in (13). As a result, the firm growth rate also drops faster in a sector with more abundant knowledge spillovers than in other sectors.

Model Implication 1: A firm's imitation rate declines with firm size while a firm's innovation rate is independent of firm size. Moreover, a firm's imitation rate drops faster in sectors with more abundant knowledge spillovers than in other sectors, which causes the cross-sector difference in the scale-independency of firm growth rate.

Notice that (12) and (13) also provide insights to the sector-specific optimal Research and development policy. Policy makers should notice that there are two types of R&D: innovation that relies on intra-firm knowledge diffusion and imitation that depends on inter-firm knowledge diffusion. Moreover, both R&D outputs have increasing return to their productivities A_M and A_N ($\frac{1}{1-\alpha}$ and $\frac{1}{1-\beta}>1$). If knowledge diffuses faster within a firm than across firms ($A_N>A_M$) and $\alpha=\beta$, increasing A_N by 1% causes a greater growth rate increment than increasing A_M by the same amount and vice versa. The reason is that firms endogenously allocate more R&D input to the type with a comparative advantage in knowledge diffusion. In order to achieve a higher economic growth rate, policy makers should tailor policies that align with the R&D type that allows for more efficient knowledge diffusion.

Polices to support imitation (push up A_M) include subsidizing cross-firm R&D cooperation, facilitating labor turnover, encouraging universities to disseminate knowledge to the public, etc. The reverse of the above policies can support innovation (push up A_N).

2.3 Determinants of Firm Size Distribution

To provide economic context for (10), I want to compare it with an AR(1) process. $R_{f,t}$ here is a random variable while for a typical AR(1) process R is a constant. In the AR(1) process $Z_{t+1}=RZ_t+L_t$, R measures persistency and L represents the randomness of the stochastic process. Similarly, in (10) $R_{f,t}$ measures the persistency of firm size, or to what extent current firm size affects future firm size by providing private knowledge capital for future innovation. $L_{f,t}$ indicates how much firms can learn from public knowledge capital which is independent of current firm size.

Imagine an economy without imitation, which means eliminating $L_{f,t}$ in (10). Starting from a sector with many equally sized firms, repeat the process for $z_{f,t+1}=R_{t+1}z_{f,t}$ for numerous periods and firms will end up with different sizes because they have different draw of "luck" in their innovation history. Overtime, firm size dispersion will grow without bound. The volatility of innovation shocks $\varepsilon_{f,t}^{n}$ determines how fast the size dispersion explodes. In the real world with chances to learn from other firms, $L_{f,t}$ constrains and attenuates the size dispersion generated by innovation shocks. In equilibrium, firm size heterogeneity is constant over time with imitation.

Proposition 1 *By theorem 5 in Kesten (1973), the firm size distribution* $\{z_{f,t}\}$ *in a given sector follows a*

Pareto distribution with scale parameter μ, such that $E\left(R_{f,t}\right)^{\frac{1}{\mu}}=1$, if $\{R_{f,t},L_{f,t}\}$ in the market size dynamics (10) are independently and identically distributed[13] *over time and across firms.*

Lemma 2 *When $\{\log\left(R_{f,t}\right)\}$ follows a normal distribution with variance σ_r^2, and $\{R_{f,t},L_{f,t}\}$ in the market size dynamics (10) are independently and identically distributed over time and across firms, there is a closed form solution for μ:*

$$\mu=1-\frac{2ln\left\{E\left(R_{f,t+1}\right)\right\}}{\sigma_r^2}\approx1+\frac{2\frac{\bar{l}}{1+\bar{r}+\bar{l}}}{\sigma_r^2}. \tag{14}$$

where $\bar{l}$ $(\bar{r})$ is the average imitation (innovation) rate in the sector. (14) highlights two offsetting forces shaping firm size distribution: the innovation shock's volatility creates firm size difference while imitation reduces the difference. As mentioned in the last paragraph, innovation shock's volatility σ_r^2 determines how fast firm size dispersion explodes without imitation. On the other side, imitation's relative contribution to the gross growth rate $\frac{\bar{l}}{1+\bar{r}+\bar{l}}$ defines the power of mean reversion to constrain the firm size dispersion from exploding. In total, firm size heterogeneity[14] $\frac{1}{\mu}$ declines with the relative magnitude between these two offsetting forces.

In addition, since abundant cross-firm knowledge spillovers or high A_M increases imitation's relative contribution to the gross growth rate $\frac{\bar{l}}{1+\bar{r}+\bar{l}}$, a sector with more abundant cross-firm knowledge spillovers should also have more homogeneous firm size distribution than other sectors.

Model Implication 2: Firm size heterogeneity declines with the relative magnitude between imitation's gross growth rate contribution and innovation risk's volatility $\frac{\frac{\bar{l}}{1+\bar{r}+\bar{l}}}{\sigma_r^2}$.

Model Implication 3: Sectors with more abundant cross-firm knowledge spillovers have more homogeneous firm size distribution than other sectors.

2.4 Growth Rate Volatility Decomposition

$$Var\left(g_{f,t}\right)=Var\left(R_{f,t}\right)+\frac{Var\left(L_{f,t}\right)}{Z_{f,t}^2} \tag{15}$$

According to the market share dynamics (10) and (15), every firm's growth rate is subject to two shocks: innovation risk and imitation risk. The relative weight of these two risks is different cross firms. For large firms, the main risk component is innovation risk; while for small firms, the major component is imitation risk. Overall, firm's growth volatility declines with its size, because innovation risk is the same across firms, while imitation risk's contribution to total volatility decreases with firm size.

Decomposing firm volatility into innovation risk and imitation risk sheds light on recent discoveries about the converging firm growth volatility among small private firms and large public firms. Comin and Mulani (2006, 2007) and Davis, Haltiwanger, Jarmin and Miranda (2006) find that U.S. large public firms' volatility rises, while small private firms' volatility declines in the last several decades. One possible mechanism to explain these two concurrent facts is that some policy or technology changes encouraged firms to invest more in innovation and less in imitation (A_N rises and/or A_M decreases). As a result, firms undertook riskier projects, when they allocate more generous fund to innovation; the opposite happens, when firms choose imitation projects with more limited fund. Such changes induce innovation volatility $Var\left(R_{f,t}\right)$ to rise and imitation volatility $Var\left(L_{f,t}\right)$ to drop at the same time. Since $Var\left(R_{f,t}\right)$ is the major risk component for large firms and $\frac{Var\left(L_{f,t}\right)}{Z_{f,t}^2}$ is the major risk

[13] The independent assumption is unnecessary according to Goldie (1991).

[14] For a *Pareto* distribution with scale parameter μ, $\frac{1}{\mu}$ is equal to the standard deviation of log scale firm size, which is commonly used as a measure of firm size heterogeneity in the literature.

component for small ones, the increment of $Var\left(R_{f,t}\right)$ dominates the decrement of $\frac{Var(L_{f,t})}{Z_{f,t}^2}$ for large firm's volatility, meanwhile the decrement in $\frac{Var(L_{f,t})}{Z_{f,t}^2}$ overweights the increment of $Var\left(R_{f,t}\right)$ for small firms.

There are existing literatures on the declining knowledge spillovers which deters imitation and encourages innovation. Caballero and Jaffe (1993) and Rosell and Agrawal (2006) find that the potency of spillover from old ideas to new knowledge generation has been declining over last century. The policy changes started from Bayh-Dole Act (35 USC 200-212) 1980, which grants patent to inventors with federal assistance. Since then US Patent Law has been amended several times to include broader and broader infringement definition. All these policy changes encourage innovation and limit the imitation. Even universities, whose traditional role was to disseminate knowledge, had became more and more commercial oriented.

Another related discovery is the divergence between moderating aggregate volatility and rising firm level volatility for public traded firms. Comin and Mulani (2006) propose an explanation also based on changing R&D activity: firms spend more resource on *Embodied* innovations and less on *Disembodied* innovations. The first type of R&D is patentable, so that firms can appropriate all the benefit it generates. The second type of R&D is hard to patent and easy to reverse engineer. The firm that develops a disembodied innovation cannot appropriate the benefits enjoyed by the other firms when adopting it. The comovement across firms weakens, when there is fewer disembodied innovations to be imitated by everyone at the same time. Since GDP volatility is the summation of individual firm's volatility and the covariance between firm's grow rates, weaker comovement reduces GDP volatility.

Summating the firm dynamics in (10) to aggregate level may provide a coherent answer to the both the volatility convergence between large and small firms at firm level and the moderating aggregate volatility at macro level. The key is firm's changing R&D pattern: less imitation and more innovation.

2.5 General Equilibrium

In general equilibrium, the marginal firm value, v; the growth rate in the number of goods, g; the average firm innovation rate, $\bar{r}$; the average firm imitation rate, $\bar{l}$; the nominal GDP, PY; and the number of firms, M_F; are solved using (16) to (21).

$$v = \frac{\frac{PY}{\sigma}}{1 - \frac{\rho}{(1+g)}\left[1 + (1-\alpha)\,\bar{r}\right]} \tag{16}$$

$$g = \bar{r} + \bar{l} \tag{17}$$

$$\bar{r} \equiv E\left(r_f\right) = A_N^{\frac{1}{1-\alpha}}\left[\frac{\rho\alpha v}{(1+g)\,M_F}\right]^{\frac{\alpha}{1-\alpha}} \tag{18}$$

$$\bar{l} \equiv E\left(l_f\right) = A_M^{\frac{1}{1-\beta}}\left[\frac{\rho\beta v}{(1+g)\,M_F}\right]^{\frac{\beta}{1-\beta}} \tag{19}$$

$$M_F = \frac{(1-\beta)\,\bar{l}\rho v}{[1-\rho]\,(1+g)\,F} \tag{20}$$

$$L = \frac{\sigma-1}{\sigma}PY + \frac{\rho v\left(\alpha\bar{r} + \beta\bar{l}\right)}{(1+g)} \tag{21}$$

In (16), the marginal firm value v increases with market size PY and innovation's relative contribution to the total growth rate $\frac{\bar{r}}{1+g}$; it decreases with elasticity of substitution σ and sector growth rate g. In (17), (18), and (19), higher R&D productivity A_N and A_M have two conflicting effects on growth rate g: first, they raise the R&D inputs for given marginal firm value v; second, they reduce marginal firm value v for given R&D inputs because emerging new products squeeze current products' market share. Overall, the former effect dominates. In (20), the total number of firms M_F increases with imitation's contribution to growth rate $\frac{\bar{l}}{1+g}$ and marginal firm value v, but decreases with entry cost F. This is because new entrants want to exploit the externality from imitation but are deterred by the entry cost. In (21), firms allocate $\frac{\sigma-1}{\sigma}PY$ workers to production and $\frac{\rho v}{(1+g)}\left(\alpha\bar{r}+\beta\bar{l}\right)$ workers to R&D. L is the size of population.

Notice that the economic growth rate, or the total number of goods growth rate, g, is independent of population size L. In (16) and (21), as L enlarges, market size PY and marginal firm value v increases proportionally. But in the mean time, the larger market also accommodates more firms as indicated in (20), which keeps $\frac{v}{M_F}$ the same. As a result, the two components of the growth rate, the innovation and imitation rates defined in (18) and (19), stay the same. In contrast to other endogenous growth models that do not incorporate the scale effect of population (Jones, 1995), this model allows policy to affect economic growth. Any policy that increases A_N or A_M will boost long-term economic growth. However, unless subsidies for innovation and imitation keep $\frac{\bar{l}}{1+g}$ ratio constant, the firm size distribution will change.

3 Testing the One-Sector Model's Implications

Before testing the three implications listed above in the one-sector model section, I introduce the data briefly.

3.1 Data

The NBER Patent Citation Data comprise detailed information on almost three million U.S. patents granted between January 1963 and December 1999, more than 16 million citations made to these patents between 1975 and 1999, and around 20,000 patent assignees, 92% of whom are non-governmental organizations. I call all the organizations as "firms" henceforth. Each patent contains highly detailed information on the innovation itself, the inventors, the assignee, etc. Moreover, patents have very wide industry and geographic coverage. The patents are classified to 42 wide SIC (Standard Industrial Classification) sectors. The percentage of U.S. patents awarded to foreign inventors has risen from about 20% in the early sixties to about 45% in the late 1990s. See Hall, Jaffe and Trajtenberg (2001) for more details.

The citation data is well suited for this paper's purpose because these citations provide detailed paper trails of intellectual interactions across firms and sectors. Aggregated by industry level, the average values of time lag, the geographic distance, and the percentage of cross-firm citation indicate the pace and abundance of knowledge diffusion in each sector. Aggregated by firm level, cross-firm citations describe the sources of knowledge in each firm's R&D process. The firm-level aggregation allows for the distinction between imitation and innovation's contributions to each firm's overall growth rate, which is critical for testing the first two implications of the one-sector model. The industry aggregation allows for the testing of the model's third implication.

Here, I use patent citation[15] to measure knowledge flow. However, citations do not represent a one-to-one mapping of direct knowledge flows. A high proportional of noise may exist, because only some citations are made by the applicant, and others by the patent examiner. Jaffe, Tratjenberg and Fogarty (2000) and Duguet and MacGarvie (2005) justify the use of aggregate patent citations as an indicator of knowledge spillovers based

[15]Patents cite other patents as "prior art," with citations describing the property rights conferred. While a patent grants the assignee the right to exclude others from practicing the invention described in the patent, it does not necessarily grant the owner the right to use the invention without the permission of cited assignees.

on a survey of patent inventors in the U.S. and firms in France. They conclude that some of the citations are associated with real knowledge flow, and patent citations aggregated at the industrial or regional level are valid measures of knowledge flow.

3.2 Implication 1: Scale-Independency of Firm Growth Rate

The one-sector model's implication 1 is: A firm's imitation rate declines with firm size, while a firm's innovation rate is independent of firm size. Moreover, a firm's imitation rate drops faster in a sector with more abundant knowledge spillovers than in other sectors, which causes the cross-sector difference in the scale-independency of firm growth rate.

In this subsection, I first demonstrate that the growth rate of a surviving firm has various scale dependencies in different sectors. I then attribute the above phenomenon to the scale-dependent imitation rate and its cross-sector differences.

Figures 3 and 4[16] show that firm growth rate in the "Petroleum and natural gas extraction and refining" sector is almost independent of firm size, while in "Office computing and accounting machinery" it drops rapidly as firm size increases. Notice that the former sector has a more heterogeneous firm size distribution than the latter. In Figure 3 firm size (growth rate) is measured by French manufacturing firms' total revenues (growth rate of total revenues) in the Amadeus Database, while in Figure 4 firm size (growth rate) is measured by a firm's number of patents (growth rate of number of patents) in the NBER Patent Citation Database. In the model, the firm growth rate in terms of number of goods or total revenues are the same.

For every sector, I run the following regression with both NBER Patent Citation Data and French manufacturing firm data:

$$g_{f,t} = a_{s,t} - b_{s,t} \ln (ps_{f,t}) .$$

where $g_{f,t}$ is firm f's growth rate at time t. $ps_{f,t}$ is the number of patents granted to firm f by the beginning of time t (or firm f's total revenue at time t in the French firm dataset). $a_{s,t}$ and $b_{s,t}$ are sector-specific. A larger $b_{s,t}$ means firm growth rate drops faster with firm size (or firm growth rate is more scale-dependent) in sector s at time t.

In Figures 5 (Figure 6), each scatter point represents one sector and the numbers label the 4 digit NAICS 2002 industry classification (SIC patent classification in the US Patent Office). The firm size dispersion measure is the standard deviation of log scale firm revenue (patent stock) in Figure 5 (Figure 6). In both graphs, b_s declines with the firm size heterogeneity measure. In other words, firm growth rate is more scale-dependent in a sector with homogeneous firm size distribution than other sectors.

This implication also predicts that when firm growth rate is broken down into innovation rate and imitation rate, the scale-independency of firm growth rate comes purely from the imitation rate (13).

A challenge in testing this implication is to estimate a firm's imitation rate $r_{f,t}$ and innovation rate $l_{f,t}$. Typically, we only observe a firm's overall growth rate, and it is difficult to tell what share is attributable to innovation based on a firm's private knowledge and what share originates from imitation based on public knowledge. The use of patent citations is a promising approach to solve this problem because they indicate the source of knowledge used during the patent invention. Within-firm citations (cross-firm citations) indicates that the citing firm uses its private (public) knowledge when creating a new patent.

At time t, firm f's patent stock growth rate $g_{f,t}$ is split into an innovation rate $r_{f,t}$ and imitation rate $l_{f,t}$,

[16] The x-axis values are discounted by sector average so that the two sectors have similar domains in firm size.

according to the ratio between within-firm citation and cross-firm citation[17].

$$\hat{g}_{f,t} = \frac{\text{No. of new patents}_{f,t}}{\text{patent stock}_{f,t}}$$

$$\hat{l}_{f,t} = \hat{g}_{f,t} \frac{\text{No. of cross-firm citations}_{f,t}}{\text{No. of total citations}_{f,t}}$$

$$\hat{r}_{f,t} = \hat{g}_{f,t} \frac{\text{No. of within-firm citations}_{f,t}}{\text{No. of total citations}_{f,t}}$$

Consider the following example. Firm f had ten patents at the beginning of year t. It invented five new patents during year t. In these five patent applications, firm f's scientists cited 30 patents held by other firms and cited its own patents 20 times. Firm f's patent stock growth rate at year t is $\hat{g}_{f,t} = \frac{5}{10} = 50\%$; the innovation rate is $\hat{r}_{f,t} = \hat{g}_{f,t} * \frac{30}{30+20} = 30\%$; and the imitation rate is $\hat{l}_{f,t} = \hat{g}_{f,t} * \frac{20}{30+20} = 20\%$.

In order to reflect the quality of the information transmitted in each citation count, I adjust the pure citation count by assigning a greater weight to a citation with shorter time lag. For example, if the citation time lag is n years, this citation is given a weight of $(1-\delta)^n$. δ is the knowledge capital depreciation rate. The three model implications are almost not affected if I let the discount rate vary between 0 and 0.9. I use $\delta = 0.1$ in the following regressions. One reason to add time discount is that citations with shorter time lag transfer more frontier knowledge on average. Another reason is that firms usually cite inside patents sooner than outside patents. Without the time discount adjustment, I underestimate the inside knowledge flow and over-estimate the outside knowledge flow.

I run the following two regressions for every sector s and time t. Again a larger $br_{s,t}$ ($bl_{s,t}$) means innovation rate (imitation rate) is more scale-dependent in sector s at time t.

$$\hat{r}_{f,t} = ar_{s,t} - br_{s,t} \ln (ps_{f,t})$$

$$\hat{l}_{f,t} = al_{s,t} - bl_{s,t} \ln (ps_{f,t})$$

Figure 7 shows the results for 1990. Similar patterns are exhibited in other years, also. Each scatter point represents one sector and the numbers label the SIC patent classification in the US Patent Office. The imitation rate scale-dependencies $\left\{\hat{b}l_{s,t}\right\}$ for all sectors are around 0.1 to 0.3 and significantly different from 0 for all sectors. In contrast, innovation scale-dependencies $\left\{\hat{b}r_{s,t}\right\}$ are around 0 and 0.05[18] which are, for most of them, not statistically significant. In addition, scale-dependency of imitation rate $\hat{b}l_{s,t}$ decreases with the sectoral firm size heterogeneity measure. This implies that the imitation rate is more scale-dependent in sectors with a more homogeneous firm size distribution. While scale-dependency of innovation rate $\hat{b}r_{s,t}$ are independent from sectoral firm size heterogeneity measure.

In summary, when the growth rate is split into an innovation rate and imitation rate, the scale dependence of the growth rate comes only from the scale dependence of the imitation rate, since the innovation rate is independent of firm size. The firm growth rate drops faster in sectors with more homogenous firm size distribution because the imitation rate reduces faster in those sectors.

[17] I only use within-sector citations made and received by U. S. firms.

[18] The outlier, sector 51, has only 20 firms applying for patents in that year. $\hat{b}r$ is not statistically significant.

3.3 Implication 2: Determinants of Firm Size Heterogeneity

The one-sector model's second implication is: Firm size heterogeneity declines with the relative magnitude between imitation's contribution to gross growth rate and innovation risk's volatility $\frac{\frac{\bar{l}}{1+\bar{r}+\bar{l}}}{\sigma_r^2}$.

For a *Pareto* distribution, the commonly used measure of firm size heterogeneity, i.e. standard deviation of log-scale firm size, is the reciprocal of the *Pareto* distribution scale parameter μ. (14) predicts that μ increases with imitation's contribution to gross growth rate, $\frac{\bar{l}}{1+\bar{r}+\bar{l}}$, and decreases with innovation shock's volatility, σ_r^2. In sector s and year t, $\bar{r}_{s,t}$ and $\bar{l}_{s,t}$ are given by the average of $\hat{r}_{f,t}$ and $\hat{l}_{f,t}$ for all firms that applied for patents in sector s and year t; $\hat{\sigma}_{rs,t}^2$ is estimated by the standard deviation of $ln\left(\frac{1+\hat{r}_{f,t}}{1+\bar{r}_{s,t}+\bar{l}_{s,t}}\right)$ in (11); sdlnps is the standard deviation of log-scale patent stock. Therefore, the model predicts that sdlnps increases in $\hat{\sigma}_{rs,t}^2$ and decreases in $\frac{\bar{l}}{1+\bar{r}+\bar{l}}$. $\frac{\bar{l}}{1+\bar{r}+\bar{l}}/\hat{\sigma}_{rs,t}^2$ is the relative magnitude of these two offsetting forces.

In Figures 8, each scatter point represents one sector and the numbers label the SIC patent classification in the US Patent Office. It shows exactly what the model predicts. In these figures, the y-axis is the firm size heterogeneity measure s.d. of log scale patent stock and the x-axis is $\frac{\bar{l}}{1+\bar{r}+\bar{l}}/\hat{\sigma}_{rs,t}^2$. Therefore, it supports the claim that when imitation's force dominates that of innovation risk's volatility, firm size distribution becomes more homogeneous.

3.4 Implication 3: Knowledge spillovers Efficiency and Firm Size Heterogeneity

The one-sector model's third implication is: Knowledge spillovers is more abundant in sectors with more homogeneous firm size distribution.

I measure knowledge spillovers efficiency by the percentage of cross-firm citations among total citations and the citation time lag of cross-firm citations. The share of cross-firm citations among all citations indicates how likely it is that knowledge spillovers crosses firm borders. Figure 9 shows that the proportion of cross-firm citations is negatively correlated with sectoral firm size heterogeneity.

The citation time lag, the interval between the application time of the citing patent and the application time of the cited patent, indicates the time needed for knowledge to travel between the inventors of these two patents. The shorter citation time lag for cross-firm citations indicates more abundant knowledge spillovers. Notice that the two inventors may take longer to exchange information if the geographic distance between them is larger. The great circle distance between the first inventor of the citing and the first inventor of the cited patent[19] measures how far knowledge travels.

Take the "Office computing and accounting machinery" and "Petroleum and natural gas extraction and refining" industries, for example. Figures 10 and 11 show that knowledge diffusion is more likely across firm borders and faster in the former industry. Notice that the former sector has a more homogeneous firm size distribution than the latter sector. The gap between these two sectors shrinks as the time lag becomes longer, but still exists even after a lag of 20 years.

The fixed-effects OLS regressions in Table 1a give the determinants of cross-firm citation time lag with U.S. citations. Since the time lag of repetitive citations overestimates the knowledge spillovers time lag, I only include a citation the first time the citing firm cites the cited patent. First-time citations account for around 70% of all citations. The regression results are similar and more significant when all citations are included.

In the first regression, there are state pair fixed effects to capture time invariant unobserved variables that may have an impact on information diffusion between the citing state and the cited state. In the second

[19]The patent inventors are required to report their mailing address. From the Census 2000 U.S. Gazetteer Files, I identify over 90% of U.S. inventors' geographic locations by their five-digit ZIP code's latitude and longitude. Using both sides' latitude and longitude data, the great circle distance between the citing patent and the cited patent is calculated by the method in Sinnott (1984).

regression, the sector fixed effects are included to take care of sector-specific time invariant elements that may affect within-sector knowledge spillovers. If sectoral firm size heterogeneity (s.d. log(patent stock)) changes over time, the model implies that citation time lag should move in the same direction. In the third regression, both types of fixed effects are considered. In the fourth column, I control the citing firm and cited firm firm-pair fixed effects. Year dummies for the citing patent application year are included in all regressions.

In all regressions, the citation time lag is longer if geographic distance is larger, the citing organization is smaller, the cited organization is smaller, the sector size is smaller, and the sectoral firm size distribution is more heterogeneous. Distance delays the exchange of knowledge because it enlarges communication cost. Larger firms are quicker to acquire information, because they are on average older and have better connections due to a more established social network. A larger industry tends to have faster knowledge diffusion because scientist density is higher. Table 1a shows that the sectoral firm size heterogeneity has the predicted positive effect on citation time lag. One standard deviation change in s.d. log(patent stock) (0.53) causes the citation time lag to increase by 0.44 (0.766*0.53) to 1.55 (2.93*0.53) years, keeping other conditions constant.

In summary, there is a greater proportion of cross-firm citations when the sectoral firm size distribution is more homogeneous. Additionally, among the cross-firm citations, citation time lag is shorter in sectors with more homogeneous firm size distribution, controlling for the sizes of the citing and cited firms, the size of the sector, and state-pair, sector and firm-pair fixed effects. These results support the third implication of the theoretical model: knowledge spillovers is more abundant in sectors with a homogeneous firm size distribution.

4 Multi-Sector Model

When organization size is measured by the number of patents, organization size distribution within each sector follows a distinct *Pareto* distribution with scale parameter ranging from 0.29 to 3[20]. When all the patenting organizations are pooled together, the organization size distribution also follows a *Pareto* distribution with a scale parameter close to 1.68. Note that one organization may apply for patents in multiple sectors in one period; the organization size in the pooled distribution of the whole economy is the summation of its size in all sectors. This result corroborates the stylized facts in Helpman, Melitz and Yeaple (2004), with firm size measured by number of employees. In their paper, every sector s has a *Pareto* firm size distribution with special scale parameter μ_s, while the aggregate economy also has a *Pareto* firm size distribution with scale parameter μ close to 1. To this point, no research has been conducted to explain the universal *Pareto* distribution of firm size in each sector and in the entire economy.

The following phenomena inspire me to consider firm size dynamics from a multi-sector perspective. First, Tables 2 and 3 show that many firms develop products in multiple sectors. Moreover, larger firms operate in more sectors. Table 2 is borrowed from Broda and Weinstein (2007)[21], which "highlights the multi-product nature of firms in these markets." It demonstrates that firms with higher sales in dollar value also sell a greater number of goods and sell in more sectors in the second to the fifth columns. Table 3 shows a similar result in patent data: organizations that own more patents also apply for patents in more patent categories.

Second, in NBER Patent Citation Data, 37% of all citations are cross-sector citations. This suggests that knowledge spillovers exists not only within but also across sectors. In Tables 4a, the row number represents the citing sector, and the column number represents the cited sector. The (i,j) element of the matrix is the percentage of citations given by sector j to sector i. There are 42 sectors in total, I only pick 6 for illustration. Every sector

[20] Estimated by French Manufacturing Firm Data from Amadues.

[21] Streitweiser (1991), Jovanovic (1993) and Bernard, Redding and Schott (2006) also found a similar extent of industry diversification in U.S. firms or plants. "UPCs" in the second column means "Universal Product Codes," commonly referred to as "barcodes." "Share" in the last column means the total market share of the firms within each group.

gives a large proportion of citations to the patents in the same sector, but also allocates a small proportion of citations to patents in every other sector. In Table 4b, I adjust the original percentage of cross-sector citation by the cited sector's weight in the dataset. The table shows that every sector cites itself over-proportionally and cites other sectors under-proportionally in most cases, but there are some sectors receive over-proportional citations (the blue cells). These blue cells inducate that the cited sector contributes above average knowledge to the citing sectors.

The multi-sector model expands upon the one-sector model by adding inter-sector knowledge spillovers. With inter-sector knowledge spillovers, a firm's growth dynamics in every sector and in the entire economy are subject to the effects from all sectors. The similarity in growth dynamics confirms that firm size distributions, whether measured within one sector or in the entire economy, all converge to the *Pareto* distribution.

Without inter-sector knowledge spillovers, firm growth dynamics in different sectors would be independent. The summation of several independent *Pareto* distributed variables is still *Pareto* distributed, but the scale parameter should be equal to the minimum of the component distribution scale parameters[22].

In contrast, the firm- or establishment-level data (Figures 1 and 2) show that the scale parameter of all firms' distribution lies between the component sectors' scale parameter values. Therefore, the firm size dynamics in different sectors must be dependent. I show in the multi-sector model that inter-sector knowledge spillovers is one of the potential forces allowing for the interactions of cross-sector dynamics.

The model is as follows. A representative firm f operates in K sectors. Firm f's size in all sectors at time t is summarized by a K dimension real vector $Z_{f,t}$. The k^{th} element of $Z_{f,t}$, $Z_{f,t}^{k}$, represents the number of products in the k^{th} industry invented by firm f. Firm f can apply its private knowledge capital in sector i, $Z_{f,t}^{i}$, to the innovation in any sector j, where $i, j \in \{1,\ 2,\ ...,\ K\}$ with production function $A_N^{ij}\left(N_{f,t}^{ij}\right)^{\alpha}\left(Z_{f,t}^{i}\right)^{1-\alpha}$. A_N^{ij} is the ability to apply sector i's knowledge to innovate in sector j (I call it ij type innovation). $N_{f,t}^{ij}$ is firm f's research hours spent in ij's type of innovation. Firm f utilizes public knowledge capital, $\bar{Z}_t$, for imitation in any sector j with production function $A_M^{j}\left(M_{f,t}^{j}\right)^{\beta}\left(\bar{Z}_t\right)^{1-\beta}$. A_M^{j} is the ability to apply public knowledge for imitation in sector j. $M_{f,t}^{j}$ is firm f's research hours spent in j's type of imitation.

Notice that the cross-sector knowledge spillovers happens both within firm borders and across firm borders. The $\left\{A_N^{ij}\right\}, i \neq j$ measures the cross-sector but within-firm-border knowledge spillovers efficiency, while $\left\{A_M^{j}\right\}$ includes both the within-sector and cross-sector types of cross-firm knowledge spillovers, because the public knowledge capital $\bar{Z}_t$ contains public knowledge from all sectors. This assumption accords with the data for cross-firm citations, where cross-firm citations occur both within and across sectors.

Firm f's manager chooses K^2+K types of R&D inputs $\left\{N_{f,t}^{ij},\ M_{f,t}^{i}\right\}$ because $i, j \in \{1,\ 2,\ ...,\ K\}$. That way, expected marginal returns from the K^2+K types of R&D are equal to their marginal costs. Solving a similar but more complex firm's problem than (2) (see Appendix B for details), the dynamics of firm size in all K sectors can be summarized by

$$Z_{f,t+1} = R_{f,t}Z_{f,t} + L_{f,t}. \tag{22}$$

$R_{f,t}$ is a $K \times K$ random matrix, and the (i, j) element $R_{f,t}^{ij}$ measures the success rate when firm f uses its private knowledge from sector j to innovate in the creation of new products in sector i. $L_{f,t}$ is a K dimensional random vector. The k^{th} element of $L_{f,t}$ is the number of imitated products in sector k which firm f has invented. For instance, the size dynamics of firm f's branch in sector k are:

$$Z_{ft+1}^{k} = R_t^{k1}Z_{ft}^{1} + ... + R_t^{kk}Z_{ft}^{k} + ... + R_t^{kK}Z_{ft}^{K} + L_t^{k}. \tag{23}$$

[22] See Jessen and Mikosch (2006) and Gabaix (2008).

Proposition 3 *If $\{Z\}$ follows the dynamics in (22) and the random matrices $R_{f,t}$ and $L_{f,t}$ in (22) satisfy the restrictions in Kesten(1973) (4.9), for any vector $x \in R^K$ and $|x| = 1$, there exist some μ such that $\{x'Z\}$ follows a Pareto distribution with parameter μ.*

The above proposition means that a *Pareto* distribution exists in any subset of the economy. For example, when studying the firm size distribution of the k^{th} sector, pick $x = (0, ...0, 1, 0, ...0)$ with the k^{th} element equal to one and all others set to zero. When studying the size distribution of all firms in the entire economy, pick $x = \frac{1}{\sqrt{K}}(1, ...1, 1, 1, ...1)$, and here the total firm size is the summation of its branches size in all K sectors.

In future work, based on an open economy version of the multi-sector model, I will advocate for favorable trade policy in sectors with more intra-sector and inter-sector knowledge spillovers. On one hand, opening a sector with more abundant intra-sector knowledge spillovers will induce a larger growth rate increment in the sector. On the other hand, opening a sector that also generates more inter-sector knowledge spillovers will cause a higher growth rate increment in the entire economy. I will show that these implications can potentially explain the difference in growth performance between East Asia and Latin America.

5 Conclusion

This paper employs knowledge spillovers to examine two questions about firm size distribution: Why is firm size heterogeneity different across sectors? and Why do firm size distributions follow dependent *Pareto* distributions in every subset of the economy?

The one-sector model answers the first question using sector-specific inter-firm knowledge spillovers efficiency. In sectors with abundant knowledge spillovers, firms invest more in imitation and less in innovation; therefore imitation contributes more substantially to the overall growth rate. Since every firm has an equal chance to learn from public knowledge, imitation has a stronger influence on smaller firms' growth rates, which leads to a declining firm growth rate with firm size. Faster catch-up of smaller firms generates a more homogeneous firm size distribution.

The one-sector model implies that knowledge spillovers is more abundant, firm growth rate declines faster with firm size, and imitation contributes more to the gross growth rate in sectors with more homogeneous firm size distribution. The model has three testable implications that are supported by NBER Patent Citation Data. The advantage of this dataset is that it keeps track of inter-firm knowledge spillovers, which allows for the measurement of the speed of knowledge diffusion and the separation of the share of the innovation rate and imitation rates in the overall growth rate of the firm.

To answer the second question, the multi-sector model improves upon the one-sector model with two additional features: firms develop products in multiple sectors and cross-sector knowledge spillovers allows for dynamics to interact across all sectors. As a result, the firm growth dynamic in any subset of the economy evolves in a pattern similar to that of the whole economy. This induces a *Pareto* firm size distribution with different scale parameters in any subset of the economy.

At the aggregate level, the micro-founded models lead to policy suggestions relevant to R&D and trade. The one-sector model suggests that R&D policy be tailored independently for every sector. In order to increase firms' growth rates, policies should be in favor of innovation (imitation) when knowledge diffuses faster within (across) firm. In an open economy version of this paper, opening sectors with more intra-sector and inter-sector knowledge spillovers fosters higher growth than liberalizing trade in other sectors.

6 Appendix A: Robustness of Sectoral Firm Size Heterogeneity

This appendix introduces a method for estimating sectoral firm size heterogeneity and the datasets used for this purpose. It then shows that sectoral firm size heterogeneity varies little when firm size is measured with different proxies. As such, sectoral firm size heterogeneity is stable over time in a specific country and also highly correlated across industrialized countries.

Helpman, Melitz and Yeaple (2004) show that there are two ways to estimate the commonly used measure of firm size heterogeneity defined by the variance of log scale firm size. For a *Pareto* distributed variable X, one method is to calculate the standard deviation of log(X) which is the reciprocal of . The other method is to estimate by OLS:

$$log(Pr(X > x)) = -\mu log(x) + c$$

and use $\frac{1}{\mu}$. Theoretically, these two methods give the same estimation. In this paper, I estimate the firm size heterogeneity measure with the second method.

The datasets used here include French manufacturing firm-level data for 1997-2005 provided by Amadeus, BUREAU van DIJK, Chilean manufacturing firm-level data for 1979-1996 provided by Chile Instituto Nacional de Estadistica, and "Industry Statistics by Employment Size" provided by the U.S. Economic Census 1997 and 2002. The French dataset has information for every firm; the Chilean dataset includes only firms with more than ten employees;[23] only the U.S. dataset gives the number of firms for ten employment size categories: 1 to 4 workers, 5 to 9 workers, 10 to 19 workers, 20 to 49 workers, 50 to 99 workers, 100 to 249 workers, 250 to 499 workers, 500 to 999 workers, 1000 to 2499 workers, and 2500 or more workers. From these numbers, I can determine the rank of the firms with 1, 5, 10, 20, 50, 100, 250, 500, 1000, 2500 employees in their six-digit NAICS industry.

To show the proxy robustness of firm size heterogeneity with different firm size proxies, it is appropriate to use the data from France and Chile because they both have more than one proxy for firm size. Number of employees, operational turnover, sales, and value added are the alternative proxies for firm size. The U.S. dataset only has the number of employees as a proxy for firm size. Sdlnl, sdlny, sdlns and sdlnva are abbreviations for standard deviation of the log(number of employees), standard deviation of log(operational turnover), standard deviation of log(sales) and standard deviation of log(value added), respectively. In the French dataset (Table 5), these four measures for 81 four-digit NAICS sectors are highly correlated, with a correlation coefficient greater than 0.9 in all years and for all combinations of variables. In the Chilean dataset (Table 6), the correlation coefficients between sdlns, sdlny and sdlnva are as high as those observed in the French dataset, but those between sdlnl and the other three variables are lower and range between 0.6 and 0.8. A possible reason for this discrepancy is that Chilean firm data are truncated, since only firms with more than ten employees are part of this dataset.

The time persistence of firm size heterogeneity appears in the U.S., French and Chilean manufacturing sectors, though these different datasets cover different time intervals. The proxy for firm size is the number of employees in all data sets. In the U.S. dataset (Figure 12), the estimations for four-digit NAICS manufacturing industries in 1997 and 2002 exhibit a tight one-to-one relationship. In France (Figure 13), the estimations for four-digit NAICS manufacturing industries in 1997 and 2005 also exhibit an almost perfect one-to-one pattern.[24]

[23] If a firm size distribution measured by the number of employees follows a *Pareto* distribution with scale parameter μ, this truncation does not affect the estimation of μ, because a *Pareto* distribution has the special feature that when the distribution is truncated from the left, the rest of the distribution on the right tail is still a *Pareto* distribution with the same scale parameter, except that the new distribution starts with a higher minimum level x_m.

[24] The outlier 3122 represents the tobacco manufacturing sector. Its heterogeneity measure drops from 3.1 in 1997 to 0.9 in 2005. There were important policy changes in this sector during this eight-year period, which might induce the significant change in firm size heterogeneity. In 2001, Brussels passed a law, soon to take effect, banning mass-media advertising of tobacco and requiring

The Chilean dataset (Figure 14) has the longest time range: 20 years. There also, the estimations for four-digit ISIC manufacturing industries in 1979 and 1996 roughly follow a one-to-one relationship. Note that Chile experienced some economic reforms during this period. The outliers typically have less than 100 establishments.

The cross-country robustness test of firm size heterogeneity is based on a comparison between French and U.S. manufacturing sectors, because they both use NAICS industry classification. There are 81 NAICS four-digit manufacturing sectors in total. The number of employees is the firm size proxy in both datasets. Figures 15 and 16 show that the correlation coefficient between sdlnl, standard deviation of log(number of employees), in the two countries is 0.74 in 1997 and 0.72 in 2002. This result corroborates a similar result in Helpman, Melitz and Yeaple (2004). They find that, although the U.S. and France have different economic policies and institutions, firm size distributions for the same industries are highly correlated across countries, with a correlation coefficient of more than 0.5.

7 Appendix B: Detailed Multi-Sector Model

The consumer's problem is:

$$U = \max_{\{x_{ik,t}\}} \int_0^\infty \rho^t \left[\sum_{k=1}^{K} s_k \log\left(Y_{k,t}\right) \right] dt,$$

$$Y_{k,t} = \left(\int_0^{I_{k,t}} x_{ik,t}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}, \; k = 1, 2, ..., K.$$

s_k is consumer preference for goods in sector k or the share of income spent in sector k.

The firm's problem is:

$$\max_{\left\{N_{f,t}^{ij},\ M_{f,t}^{i}\right\},\ i,\ j \in \{1,2,...,K\}} V(z_{f,t}) = \sum_i \frac{P_{i,t} Y_{i,t}}{\sigma_i} \frac{z_{f,t}^i}{I_{i,t}} + \rho E[V(z_{f,t+1})] - \frac{\sum_{ij} N_{f,t}^{ij} + \sum_i M_{f,t}^i}{\bar{Z}_t}$$

subject to

$$z_{f,t+1} = z_{f,t} + \Delta z_{f,t}^N + \Delta z_{f,t}^M \tag{24}$$

$$\frac{\Delta z_{f,t}^{N,i}}{z_{f,t}^i} = \sum_j \left[\frac{A_N^{ij} \left(N_{f,t}^{ij}\right)^\alpha \left(z_{f,t}^j\right)^{1-\alpha}}{z_{f,t}^i} + \varepsilon_{f,t}^{N,ij} \right], \; i,\; j \in \{1, 2, ..., K\} \tag{25}$$

$$\frac{\Delta Z_{f,t}^{M,i}}{\bar{Z}_t} = \frac{A_M^i \left(M_{f,t}^i\right)^\beta \left(\bar{Z}_t\right)^{1-\beta}}{\bar{Z}_t} + \varepsilon_{f,t}^{M,i}, \; i \in \{1, 2, ..., K\} \tag{26}$$

where $\Delta z_{f,t}^N$ and $\Delta z_{f,t}^M$ are K dimension vectors. The i^{th} element of $\Delta z_{f,t}^N$ ($\Delta z_{f,t}^M$), $\Delta z_{f,t}^{N,i}$ ($\Delta z_{f,t}^{M,i}$) is the number of innovated (imitated) new goods in sector i by firm f at time t. $\left\{\varepsilon_{f,t}^{N,ij}, \varepsilon_{f,t}^{M,i}\right\}$ are i.i.d. across firms and time.

large warning labels on cigarette packages. To discourage potential new smokers, governments throughout Europe increased their cigarette taxes in 2003.

An educated guess for the firm value function is $V(z^i_{f,t}) = \sum_i v_i \frac{z^i_{f,t}}{I_{i,t}} + u$. The first order conditions and Bellman equation can be written as:

$$N^{ij}_{f,t} = \left(\frac{A^{ij}_N \alpha v_i \rho I_t}{I_{t+1} M_F}\right)^{\frac{1}{1-\alpha}} z^j_{f,t},\ i,\ j \in \{1, 2, ..., K\} \tag{27}$$

$$M^i_{f,t} = \left(\frac{A^i_M \beta v_i \rho I_t}{I_{t+1} M_F}\right)^{\frac{1}{1-\beta}} \bar{Z}_t,\ i \in \{1, 2, ..., K\} \tag{28}$$

$$v_j = \frac{P_{j,t} Y_{j,t}}{\sigma_i} + \frac{\rho I_t}{I_{t+1}}\left[1 + (1-\alpha)\sum_i \frac{I_j}{I_i} v_i A^{ij}_N \left(N^{ij}_{f,t}\right)^{\alpha} \left(z^j_{f,t}\right)^{-\alpha}\right],\ i,\ j \in \{1, 2, ..., K\} \tag{29}$$

In (27) and (28), the input in each types of R&D is proportional to the knowledge capital input. In (29), the marginal value of one product in sector j, v_j ,depends on its current profit in sector j plus its contribution to future innovation in all K sectors.

The firm size dynamic in sector i is:

$$\begin{aligned} z^i_{f,t+1} \quad = \quad & \frac{I_t}{I_{t+1}}\left(1 + A_N^{ii\frac{1}{1-\alpha}}\left[\frac{\alpha v_i \rho I_t}{I_{t+1} M_F}\right]^{\frac{\alpha}{1-\alpha}} + \varepsilon^{N,ii}_{f,t}\right) z^i_{f,t} + \\ & \frac{I_t}{I_{t+1}} \sum_{j \neq i} \left(A_N^{ij\frac{1}{1-\alpha}}\left[\frac{\alpha v_i \rho I_t}{I_{t+1} M_F}\right]^{\frac{\alpha}{1-\alpha}} + \varepsilon^{N,ij}_{f,t}\right) z^j_{f,t} + \frac{I_t}{I_{t+1}}\left(A_M^{j\frac{1}{1-\beta}}\left[\frac{\beta v_i \rho I_t}{I_{t+1} M_F}\right]^{\frac{\beta}{1-\beta}} + \varepsilon^{M,i}_{f,t}\right). \end{aligned}$$

The firm size dynamics in all K sectors are summarized by

$$z_{f,t+1} = R_{f,t} z_{f,t} + L_{f,t}, \tag{30}$$

where

$$R^{ij}_{f,t+1} \equiv \frac{I_t}{I_{t+1}}\left(\left(A^{ij}_N\right)^{\frac{1}{1-\alpha}}\left[\frac{\alpha v_i \rho I_t}{I_{t+1} M_F}\right]^{\frac{\alpha}{1-\alpha}}\right) + \varepsilon^{N,ij}_{f,t}, j \neq i,\ i,\ j \in \{1, 2, ..., K\}$$

$$R^{ii}_{f,t+1} \equiv \frac{I_t}{I_{t+1}}\left(1 + \left(A^{ii}_N\right)^{\frac{1}{1-\alpha}}\left[\frac{\alpha v_i \rho I_t}{I_{t+1} M_F}\right]^{\frac{\alpha}{1-\alpha}}\right) + \varepsilon^{N,ii}_{f,t},\ i \in \{1, 2, ..., K\}$$

$$L^i_{f,t+1} \equiv \frac{I_t}{I_{t+1}}\left(A^i_M\right)^{\frac{1}{1-\beta}}\left[\frac{\beta v \rho I_t}{I_{t+1} M_F}\right]^{\frac{\beta}{1-\beta}} + \varepsilon^{M,i}_{f,t}.$$

7.1 General Equilibrium

In general equilibrium, the marginal value of a firm, $\{v_i\}$; the innovation rate $\{\bar{r}_i\}$ and imitation rate $\{\bar{l}_i\}$; the relative number of goods $\left\{\frac{I_j}{I_i}\right\}$ $i,\ j = 1, 2, ..., K$; the number of goods growth rate g; nominal GDP PY; and number of firms M_F are solved by the following equations:

$$v_j = \frac{S_i PY}{\sigma_i} + \frac{\rho}{(1+g)}\left[1 + (1-\alpha)\left(\frac{\alpha\rho}{(1+g) M_F}\right)^{\frac{\alpha}{1-\alpha}} \sum_{i=1}^{K}\left(\frac{I_j}{I_i} v_i A^{ij}_N\right)^{\frac{1}{1-\alpha}}\right],\ i,\ j \in \{1, 2, ..., K\}$$

$$\bar{r}_i \equiv \left[\frac{\rho\alpha v_i}{(1+g)\,M_F}\right]^{\frac{\alpha}{1-\alpha}} \sum_{j=1}^{K} \left(A_N^{ij}\right)^{\frac{1}{1-\alpha}} \frac{I_j}{I_i},\ i,\ j \in \{1,2,...,K\} \tag{31}$$

$$\bar{l}_i \equiv \left(A_M^i\right)^{\frac{1}{1-\beta}} \left[\frac{\rho\beta v_i}{(1+g)\,M_F}\right]^{\frac{\beta}{1-\beta}} \frac{\sum_{j=1}^{K} I_j}{I_i},\ i,\ j \in \{1,2,...,K\} \tag{32}$$

$$g_i = \bar{r}_i + \bar{l}_i,\ i \in \{1,2,...,K\}$$

$$g_i = g,\ i \in \{1,2,...,K\}$$

$$M_F = \frac{(1-\beta)\sum_{i=1}^{K} \bar{l}_i \rho v_i}{[1-\rho]\,(1+g)\,F}$$

$$L = \frac{\sigma-1}{\sigma} PY + \frac{\rho v \sum_{i=1}^{K}\left(\alpha\bar{r}_i + \beta\bar{l}_i\right)}{(1+g)}$$

Notice that the number of goods in every sector is growing at the same speed because inter-sector knowledge spillovers keeps all sectors on the same growing track. If one sector i had been growing more slowly than other sectors for a lengthy period, its number of goods would be very small relative to other sectors. The cross-sector knowledge spillovers would push up g_i to infinity through a very large ratio, $\frac{I_j}{I_i}$, in (31) and (32), until g_i is equal to the common growth rate.

7.2 Simulation

Kesten (1973) does not provide the close-form solution for the Pareto distribution scale parameter μ in the multiple variable case. To exam how the multi-sector model fits the firm size distribution data, I estimate the distribution of $\{R_{f,t}\}$ and $\{L_{f,t}\}$ and simulate the firm size dynamics process in (30). There are 42 sectors in total, i. e. $K = 42$. I take the time period 1987 to 1997 from the NBER Patent Data to abtain the most observations each year. Elements in $\{R_{f,t}\}$ and $\{L_{f,t}\}$ are estimated as follows.

$$\hat{R}_{f,t}^{ij} = \frac{\triangle ps_{i,f,t}}{z_{j,f,t}} \frac{\text{No. of within-firm citations made by firm } f \text{ from industry } i \text{ to industry } j}{\text{No. of citations made by firm } f \text{ in industry } i},\ j \neq i,\ i, j = 1,2,...,K$$

$$\hat{R}_{f,t}^{ii} = 1 + \frac{\triangle ps_{i,f,t}}{z_{i,f,t}} \frac{\text{No. of within-firm citations made by firm } f \text{ from industry } i \text{ to industry } i}{\text{No. of citations made by firm } f \text{ in industry } i},\ i = 1,2,...,K$$

$$\hat{L}_{f,t}^{i} = \triangle ps_{i,f,t} \frac{\text{No. of cross-firm citations made by firm } f \text{ from industry } i}{\text{No. of citations made by firm } f \text{ in industry } i},\ i = 1,2,...,K$$

I fit each above element into lognormal distribution and estimate the correspondent variance and mean for the lognormal distribution. I then simulate the firm size dynamics process in (30) for 100 periods. at the end of the 100^{th} period, I record the scale parameter μ_s for each sector s and μ for the whole economy. After repeating the same simulation for 100 times, I report the mean scale parameter of the 100 simulations for each sector and for the whole economy. I plot the simulated firm size distributions for the 100^{th} simulation in Figure . The correlation between the real and mean simulated scale parameters are .

8 Appendix C: Robustness Checks

This appendix provides robustness check to the one-sector model's implications with alternative citation datasets. The first robustness check is done with random simulated citations. The second robustness check is done with all G7 country citations, which include more than 90% of all patents in NBER Patent Database, while US patents only account for about 50% of all patents.

8.1 Random Citation Data

A doubt to the one-sector model's third implication is that cross-sector citation should be fewer in a sector with more heterogeneous firm size distribution, even when cross-firm knowledge diffusion is equally complete and instant in each sector. The reason is that there are more big firms in a heterogeneous sector and a big firm is more likely to cite its own patent simply because it has more patent stock available to be cited.

I simulate such random citation datasets to mimics the environment where information is complete in every sector and compare them with the real citation data. Then I show that the real citation dataset exhibits significantly larger cross-sector differences in knowledge diffusion than the random citation datasets.

In the random citation dataset, the citing patent is kept the same as in the real citation data, but the cited patent is randomly assigned. Every existing patent in the same sector has an equal chance to be cited, regardless of the distance and other characteristics of the citing firm and the cited firm. I simulated 100 such random citation datasets. The values reported in Figure 17 are the median of these 100 datasets' results.

Figure 17 displays that cross-firm citations account for 95% to 99.9% among total random citations across sectors. Although it seems that a sector with heterogeneous firm size distribution has a less percentage of cross-firm citations than other sectors, the 5% cross-sector gap is trivial as compared with the 40% gap in the real citation dataset (see Figure 10).

I run the same regressions as those in Table 1a using the 100 random citation data sets and report the results in Table 7a. The coefficients and the robust standard errors reported are the median value of the 100 regression results. Compared with Table 1a, the regression results using random citation data show that distance does not delay knowledge diffusion as half as it does in real citation data; bigger firms do not cite outside patents faster than smaller firms at all; and citation time lag is slightly positively correlated with sectoral firm size heterogeneity, but the regression coefficients are much smaller than those in Table 1a. Notice that two factors still affect citation time lag in a similar magnitude as they do in Table 1a. The citation time lag is smaller when the cited patent is owned by a bigger firm and when the sector has a larger patent stock.

Notice that these "random citations" are not purely random, because the knowledge receiver, the citing firm, is still the same as in the real citation data, only the knowledge giver, the cited firm, is random. That is why cross-sector differences in knowledge spillovers do not disappear completely in the simulated random citation datasets.

Above all, the cross-sector differences in knowledge diffusion that exist in the real citation dataset are dramatically smaller in random citation datasets, where knowledge spillovers is equally complete and instant in all sectors by construction.

8.2 G7 Country Citation Data

The G7[25] country citation dataset also supports the one-sector model's implications. All estimation methods used are the same as those in section 3.

With similar results as Figure 7, Figure 18 and 19 show that innovation rate is independent of firm size and imitation rate declines with firm size in the G7 country citation dataset. Moreover, the scale-independency of imitation rate is negatively related to sectoral firm size heterogeneity.

In line with Figure 9, Figure 20 supports that the sectoral firm size heterogeneity is negatively related to the ratio between the imitation's contribution to gross growth rate and the innovation risk's volatility in the larger dataset with G7 country citation data.

Figure 21 shows the same negative relation between the cross-firm knowledge diffusion abundance and the sectoral firm size heterogeneity as Figure 10.

I run the same regressions as those in Table 1a using the G7 citation data and report the results in Table 8a. The sectoral firm size heterogeneity is statistically significantly positively correlated with citation time lag in the regression No. 3 with country pair-industry fixed effect; but I find the same coefficient is not statistically significant in the regression No.2 and barely significant at 10% level in the regression No. 1. My explanation is that international citations may involve more country pair-industry specific unobserved variables that are correlated with the sectoral firm size heterogeneity measure and the citation time lag at the same time, therefore firm size heterogeneity measure is not significant in the first two regressions.

9 Appendix D

In this section, I extend the basic one-sector model to allow firms combine private and public knowledge in imitation, while firms still use only private knowledge in innovation. Everything else keeps the same, except that imitated new goods production function becomes

$$E\left(\Delta z_{f,t}^{N}\right) = A_M M_{f,t}^{\beta}\left(\gamma z_{f,t} + (1-\gamma)\,\bar{Z}_t\right)^{(1-\beta)}$$

, where γ is private knowledge's share in the combined knowledge. This imitation function implies that a firm's past R&D experience helps absorb current public knowledge. In another story, positive sorting in firm's social network means what a firm expects to learn from the public increases with its own size, because firms of similar size are more likely to be connected. γ reflects the significance of positive sorting in the social network. These ideas are in line with the facts in patent citation data: when citing other firms' patents, firms with more patent stock tend to cite newer existing patents, cite larger firms' patents, and cite more diversified sources than smaller firms.

Firm's problem becomes

$$\max_{N_{f,t},\ M_{f,t}} \ V(z_{f,t}) = \frac{P_t Y_t}{\sigma}\frac{z_{f,t}}{I_t} + \rho E[V(z_{f,t+1})] - \frac{N_{f,t}+M_{f,t}}{\bar{Z}_t}$$

subject to

$$z_{f,t+1} = z_{f,t} + \Delta z_{f,t}^{N} + \Delta z_{f,t}^{M}$$

[25] Canada, France, Germany, Italy, Japan, U.K. and U.S.

$$\frac{\Delta z_{f,t}^{N}}{z_{f,t}}=\frac{A_N N_{f,t}^{\alpha} z_{f,t}^{1-\alpha}}{z_{f,t}}+\varepsilon_{f,t}^{n}$$

$$\frac{\Delta z_{f,t}^{M}}{\bar{Z}_t}=\frac{A_M M_{f,t}^{\beta}\left(\gamma z_{f,t}+(1-\gamma)\bar{Z}_t\right)^{(1-\beta)}}{\bar{Z}_t}+\varepsilon_{f,t}^{m}$$

$$N_{f,t}=\left(\frac{A_N \alpha v \rho I_t}{I_{t+1}M_F}\right)^{\frac{1}{1-\alpha}} z_{f,t}$$

$$M_{f,t}=\left(\frac{A_M \beta (1-\gamma) v \rho I_t}{I_{t+1}M_F}\right)^{\frac{1}{1-\beta}} \left(\gamma z_{f,t}+(1-\gamma)\bar{Z}_t\right)$$

$$v=\frac{P_t Y_t}{\sigma}+\frac{\rho v I_t}{I_{t+1}}\left[1+(1-\alpha)A_N\left(\frac{A_N \alpha v \rho I_t}{I_{t+1}M_F}\right)^{\frac{\alpha}{1-\alpha}}+\gamma(1-\beta)A_M\left(\frac{A_M \beta (1-\gamma) v \rho I_t}{I_{t+1}M_F}\right)^{\frac{\beta}{1-\beta}}\right]$$

Larger private knowledge's share in imitation γ boosts marginal firm value v, because future return on imitation also relies on current firm size. In the social network story, a larger size today wins the firm a better peer to imitate in the future.

Higher γ also induces larger firm size heterogeneity in the sector. When private knowledge is more important in imitation or social network is more positively assorted, sectoral firm size heterogeneity is larger for given productivity of innovation and imitation (A_N and A_M). In other words, rising γ incurs the same impact on firm size heterogeneity as rising innovation productivity A_N or decreasing imitation productivity A_M.

References

[1] Agrawal, Ajay, Devesh Kapur and John McHale. 2008. "How Do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data." Journal of Urban Economics forthcoming.

[2] Agrawal, Ajay, Iain Cockburn and John McHale. 2006. "Gone But Not Forgotten: Knowledge Flows, Labor Mobility, and Enduring Social Relationships." Journal of Economic Geography 6:571-91.

[3] Albuquerque, Rui and Hugo A. Hopenhayn. 2004. "Optimal Lending Contracts and Firm Dynamics." Review of Economic Studies 71: 285-315.

[4] Axtell, Robert L. 2001. "Zipf Distribution of U.S. Firm Sizes." Science 293:1818-20.

[5] Bernard, Andrew B., Stephen J. Redding and Peter K. Schott. 2006. "Multi-product Firms and Product Switching." NBER Working Paper No. 12293.

[6] Broda, Christian and David E. Weinstein. 2007. "Product Creation and Destruction: Evidence and Price Implications." NBER Working Paper No. 13041.

[7] Cabral, Luis M. B. and Jose Mata. 2003. "On the Evolution of the Firm Size Distribution: Facts and Theory." American Economic Review 93: 1075-1090.

[8] Clauset, Aaron, Cosma Rohilla Shalizi and Mark. E. J. Newman. 2007. "Power-Law Distributions in Empirical Data." E-print arXiv:0706.1062.

[9] Clementi, Gian Luca and Hugo A. Hopenhayn. 2006. "A Theory of Financing Constraints and Firm Dynamics." Quaterly Journal of Economics 121: 229-265.

[10] Comin, Diego and Sunil Mulani. 2006. "Diverging Trends in Aggregate and Firm Volatility." Review of Economics and Statistics 88:374-83.

[11] Comin, Diego and Thomas Philippon. 2005. "The Rise in Firm-Level Volatility: Causes and Consequences." Mark Gertler and Kenneth Rogoff, eds., NBER Macroeconomics Annual 2005.

[12] Conley, Timothy and Bill Dupor. 2003. "A Spatial Analysis of Sectoral Complementarity," Journal of Political Economy, 111:311-52.

[13] Cooley, Thomas F. and Vincenzo Quadini. 2001. "Financial Markets and Firm Dynamics." American Economic Review 91:1286-1310.

[14] Cowan, Robin and Nicolas Jonard. 2004. "Network Structure and the Diffusion of Knowledge." Journal of Economic Dynamics and Control 28:1557-75.

[15] Crespi, Gustavo, Chiara Criscuolo, Jonathan E. Haskel and Matthew Slaughter. 2008. "Productivity Growth, Knowledge Flows, and Spillovers." NBER Working Papers No. 13959.

[16] Davis, Steven, John Haltiwanger, Ron Jarmin and Javier Miranda. 2006. "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms." Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds., NBER Macroeconomics Annual 2006.

[17] Dixit, Avinash K. and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." American Economic Review 67:297-308.

[18] Duguet, Emmanuel and Megan MacGarvie. 2005. "How Well Do Patent Citations Measure Flows of Technology? Evidence from French Innovation Surveys." Economics of Innovation and New Technology 14:375-93.

[19] Dunne, Timothy, Mark J. Roberts and Larry Samuelson. 1989. "The Growth and Failure of U.S. Manufacturing Plants." Quarterly Journal of Economics 104:671-98.

[20] Evans, David S. 1987. "The Relationship between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries." Journal of Industrial Economics 35:567-81.

[21] Ericson, Richard and Ariel Pakes. 1995. "Marcov Perfect Industry Dynamics: A Framework for Empirical Analysis." Review of Economic Studies 62:53-82.

[22] Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation." Quarterly Journal of Economics 114:739-67.

[23] Gabaix, Xavier. 2008. "Power Laws in Economics and Finance." NBER Working Paper No. 14299.

[24] Grossman, Gene M. and Elhanan Helpman. 1991. "Quality Ladders in the Theory of Growth." Review of Economic Studies 58:43-61.

[25] Hall, Bronwyn H. 1987. "The Relationship between Firm Size and Firm Growth in the US Manufacturing Sector." Journal of Industrial Economics 35:583-605.

[26] Hall, Bronwyn H., Adam B. Jaffe and Manuel Trajtenberg. 2001. "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper 8498.

[27] Helpman, Elhanan, Marc J. Melitz and Stephen R. Yeaple. 2004. "Export versus FDI with Heterogenous Firms." American Economic Review 94:300-16.

[28] Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." Econometrica 60:1127-50.

[29] Hsu, Wen-Tai. 2007. "Central Place Theory and Zipf's Law." University of Minnesota Working Paper.

[30] Ijiri, Yuji and Herbert A. Simon. 1964. "Business Firm Growth and Size." American Economic Review 54:77-89.

[31] Jackson, Matthew and Leeat Yariv. 2007. "Diffusion of Behavior and Equilibrium Properties in Network Games." American Economic Review 97:92-98.

[32] Jaffe, Adam B. 1986. "Technology Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value." American Economic Review 76:984-1001.

[33] Jaffe, Adam B., Manuel Trajtenberg and Michael S. Fogarty. 2000. "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors." American Economic Review 90:215-18.

[34] Jessen, Anders Hedegaard and Thomas Mikosch. 2006. "Regularly Varying Functions." Publications de L'institut Mathematique 94:171-92.

[35] Jones, Charles I. 1995. "R & D-Based Models of Economic Growth." The Journal of Political Economy 103:759-784.

[36] Jovanovic, Boyan. 2003. "The Diversification of Production." Brookings Papers on Economic Activity. Microeconomics 1993:197-247.

[37] Kesten, Harry. 1973. "Random Difference Equations and Renewal Theory for Products of Random Matrices." Acta Mathematica 131:207-48.

[38] Klepper, Steven and Peter Thompson. 2007. "Submarkets and the Evolution of Market Structure." Rand Journal of Economics 34:862-88.

[39] Klette, Tor Jakob and Arvid Raknerud. 2002. "How and Why Do Firms Differ?" Mimeo, Oslo University, Department of Economics.

[40] Klette, Tor Jakob and Samuel Kortum. 2004. "Innovating Firms and Aggregate Innovation." Journal of Political Economy 112:986-1018.

[41] Lucas, Robert E. 1978. "On the Size-Distribution of Business Firms." Bell Journal of Economics 9:508-23.

[42] Luttmer, Erzo G. J. 2007. "Selection, Growth, and the Size Distribution of Firms." Quarterly Journal of Economics 122:1103-44.

[43] Luttmer, Erzo G. J. 2008. "On the Mechanics of Firm Growth." Federal Reserve Bank Minneapolis Working Paper No. 657.

[44] Melitz, Marc J. 2003. "The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations." Econometrica 71:1695-1725.

[45] Newman, Mark E. J. 2003. "The Structure and Function of Complex Networks." SIAM Review 45:167-256.

[46] Reed, William J. 2001. “The *Pareto*, Zipf and Other Power Laws.” Economics Letters 74:15-19.

[47] Rossi-Hansberg, Esteban and Mark L. J. Wright. 2007. "Establishment Size Dynamics in the Aggregate Economy." American Economic Review 97:1639-66.

[48] Segerstrom, Paul S. 1991. "Innovation, Imitation, and Economic Growth." Journal of Political Economy 99:807-27.

[49] Seker, Murat. 2007. "A Structural Model of Establishment and Industry Evolution: Evidence from Chile." University of Minnesota Job Market Paper.

[50] Simon, Herbert A. and Charles P. Bonini. 1958. “The Size Distribution of Business Firms.” American Economic Review 48:607-17.

[51] Sinnott, R. W. 1984. "Virtues of the Haversine." Sky and Telescope 68:159.

[52] Streitweiser, Mary L. 1991. "The Extent and Nature of Establishment Level Diversification in Sixteen US Industries." Journal of Law and Economics 34:503-34.

[53] Sutton, John. 1997. “Gibrat’s Legacy.” Journal of Economics Literature 35:40-59.

[54] Sutton, John. 2002. “The Variance of Firm Growth Rates: The Scaling Puzzle.” Physica A 312:577–90.

[55] Sutton, John. 2007. "Market Share Dynamics and the ’Persistence of Leadership’ Debate." American Economic Review 97:222-41.

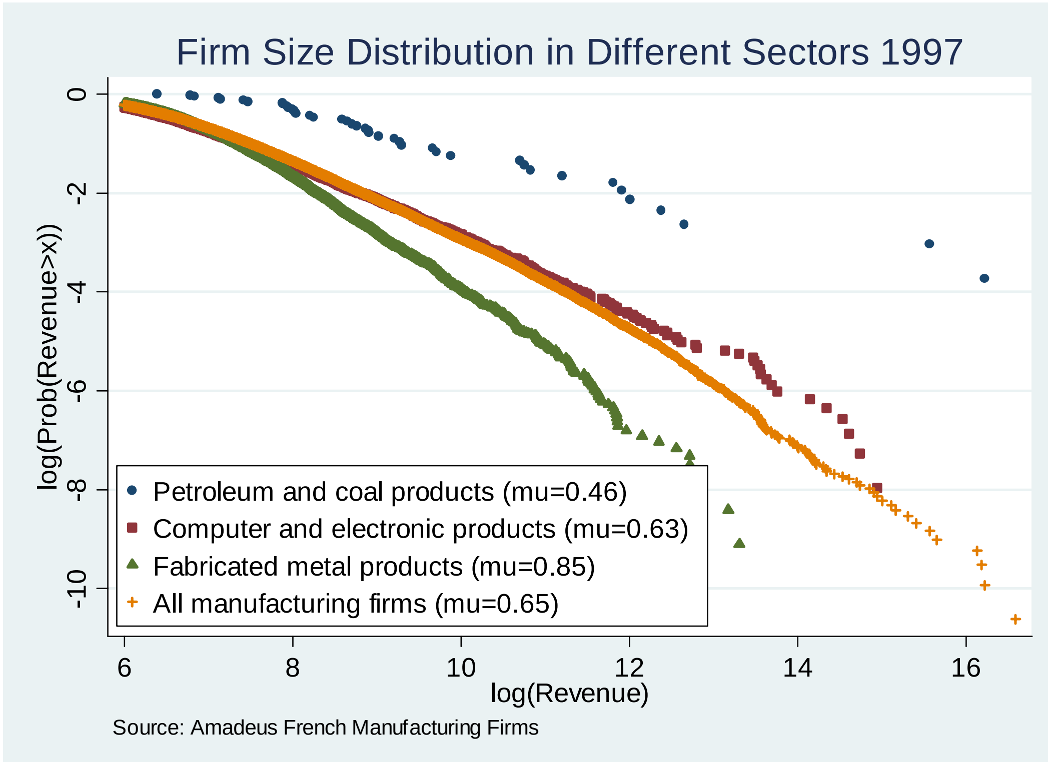

Figure 1

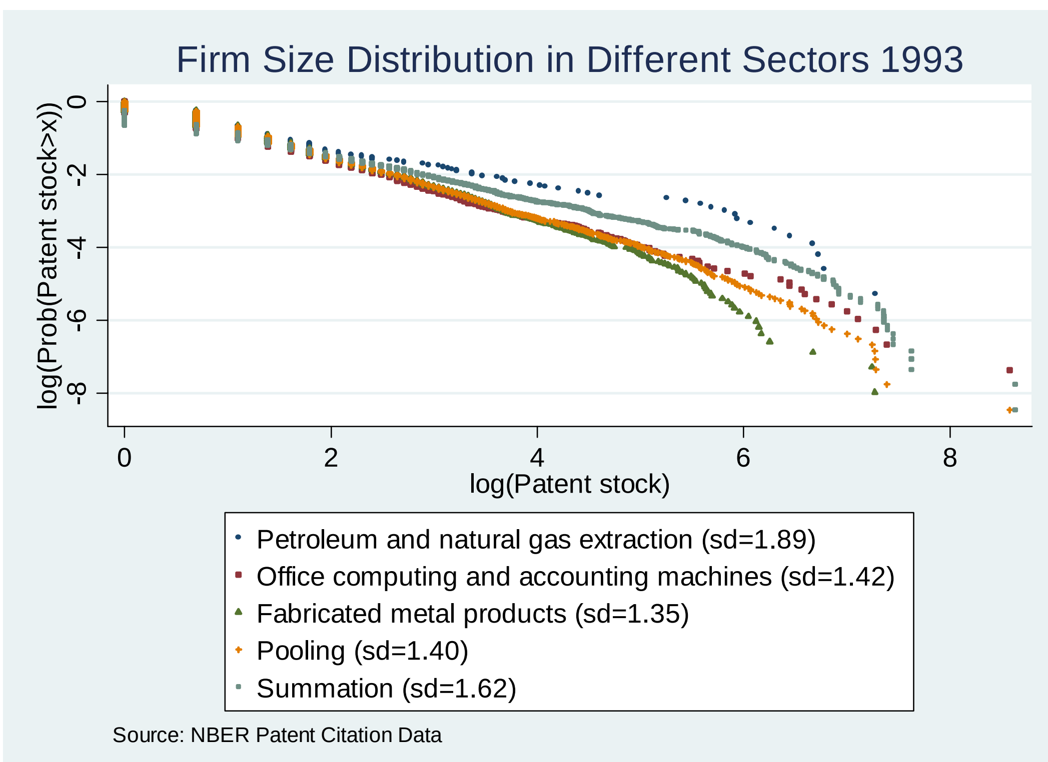

Figure 2

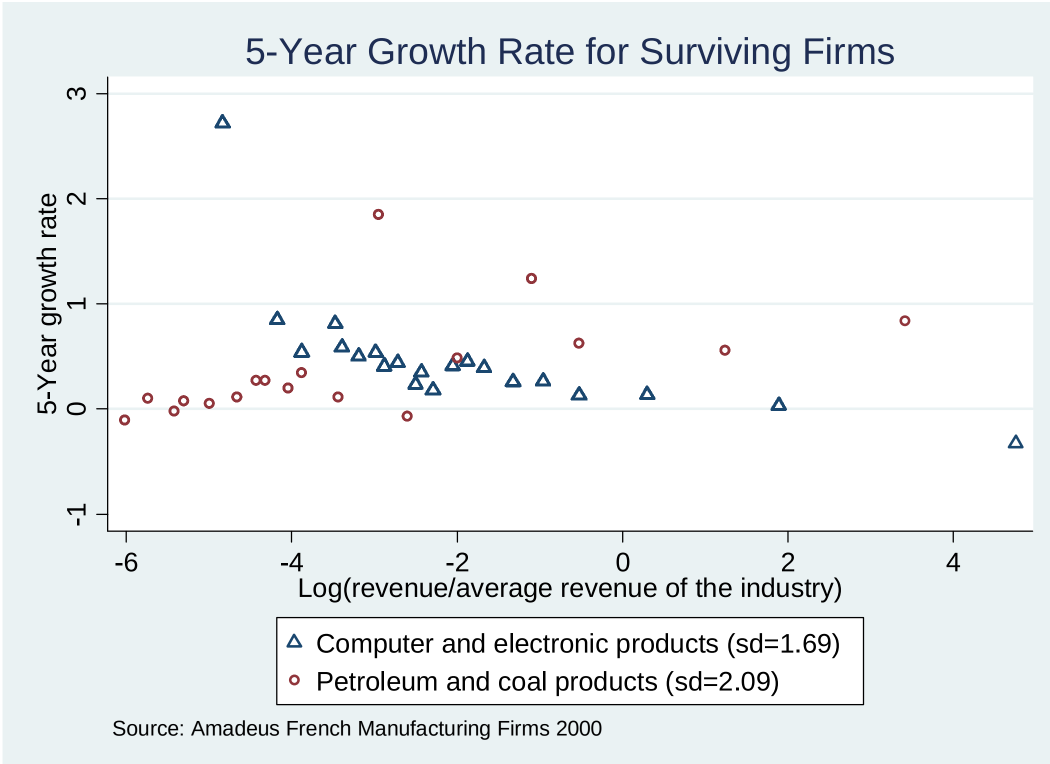

Figure 3

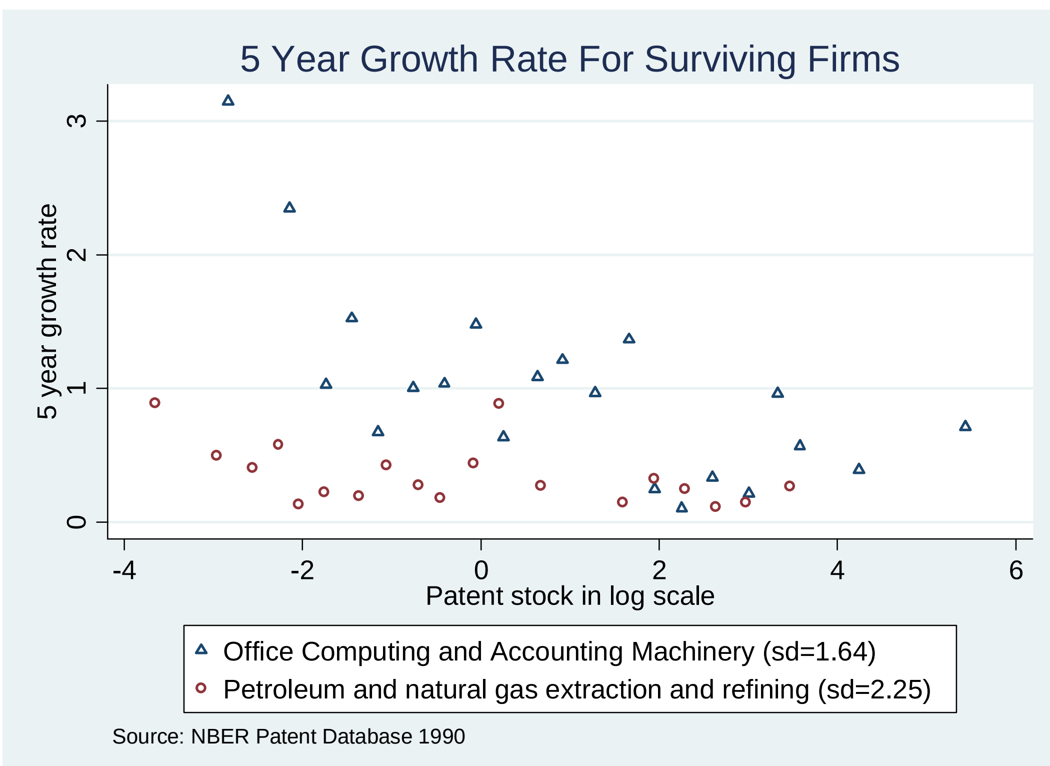

Figure 4

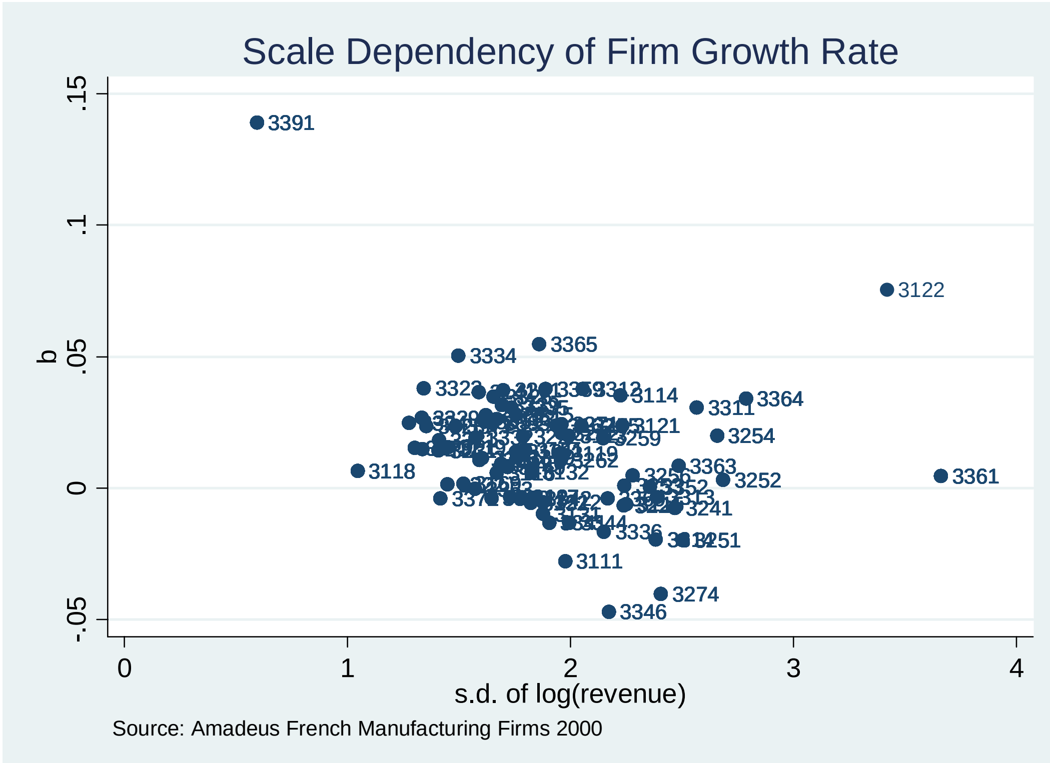

Figure 5

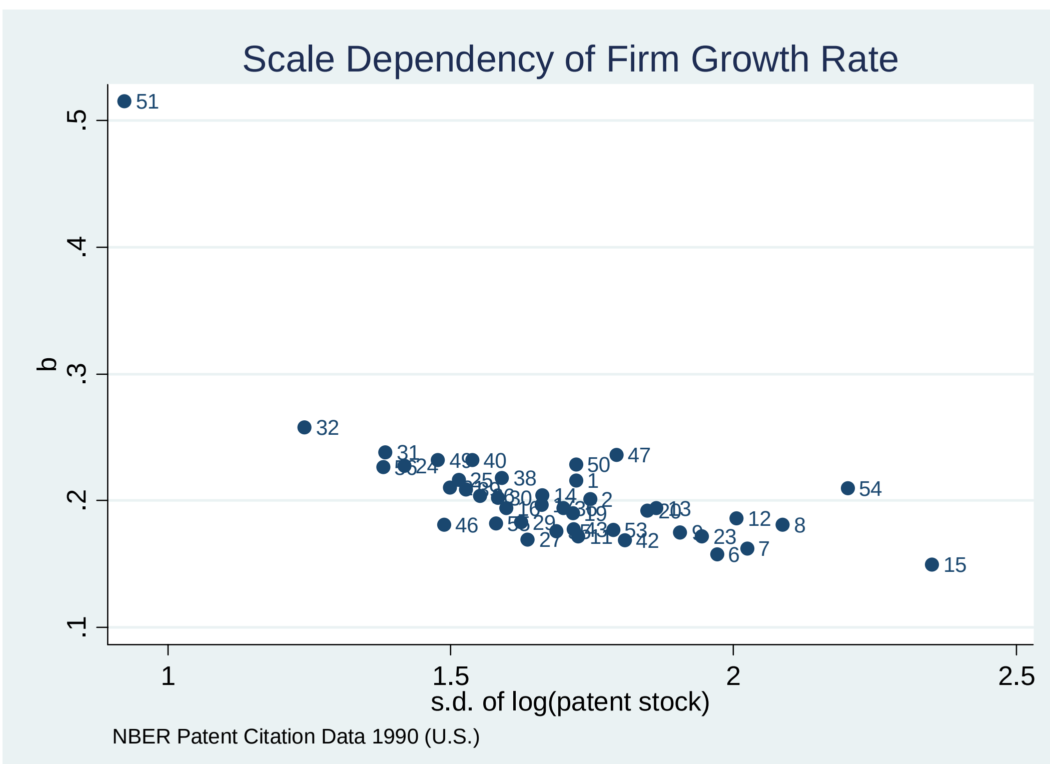

Figure 6

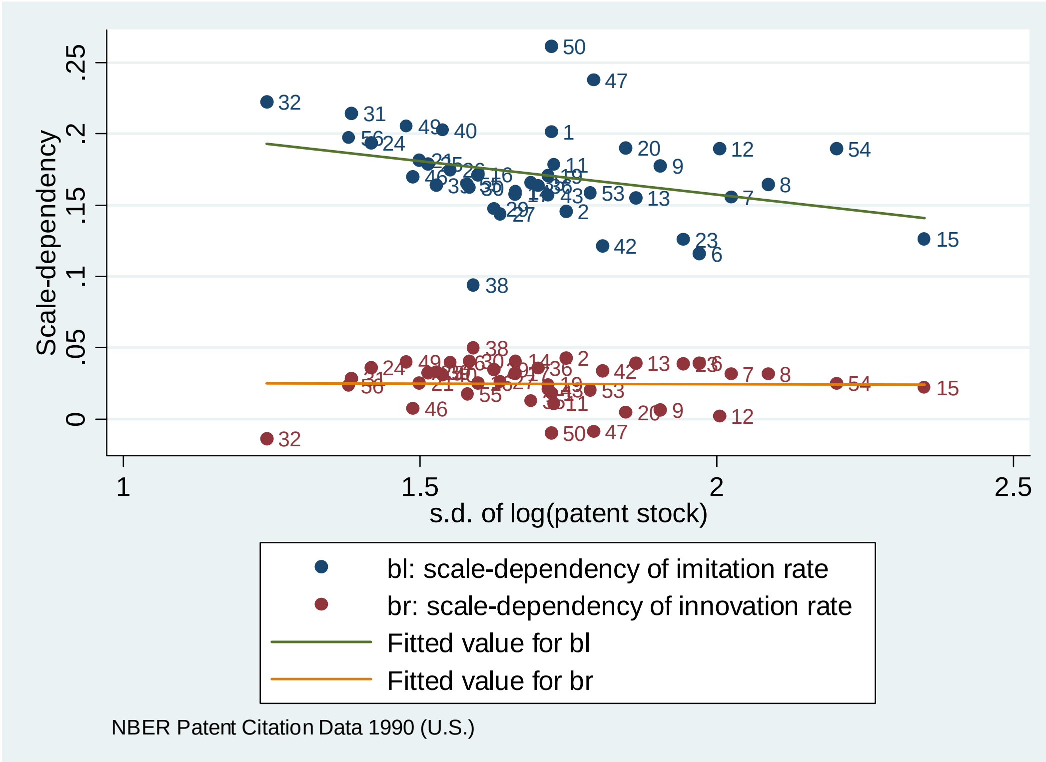

Figure 7

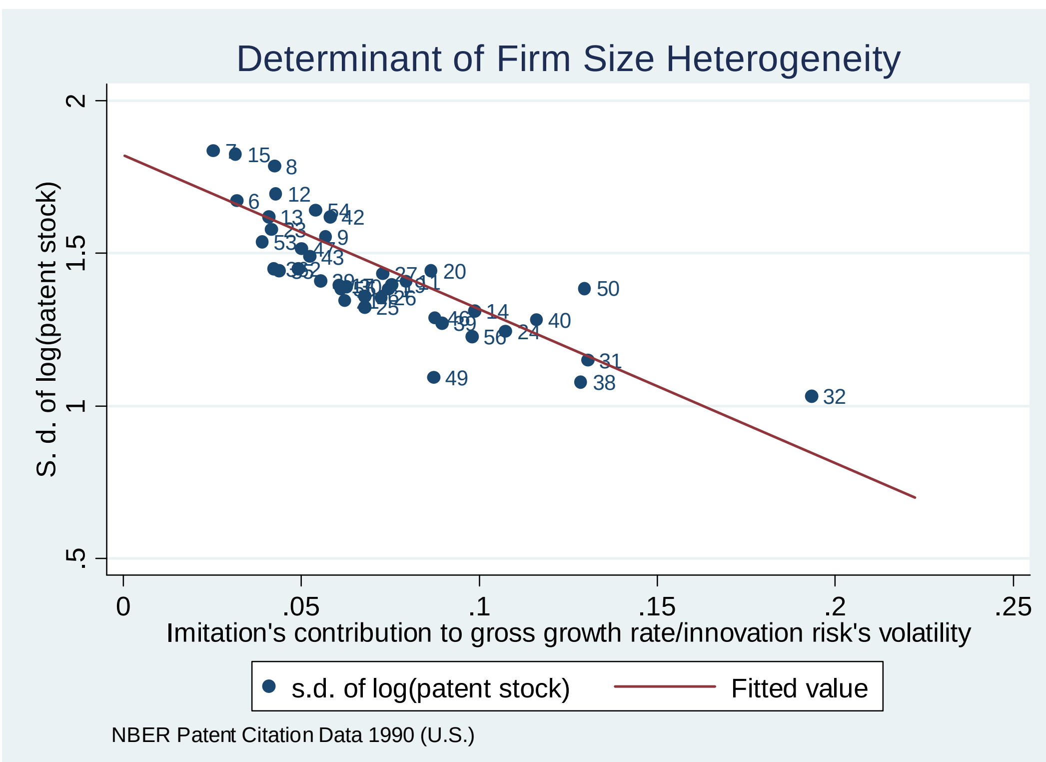

Figure 8

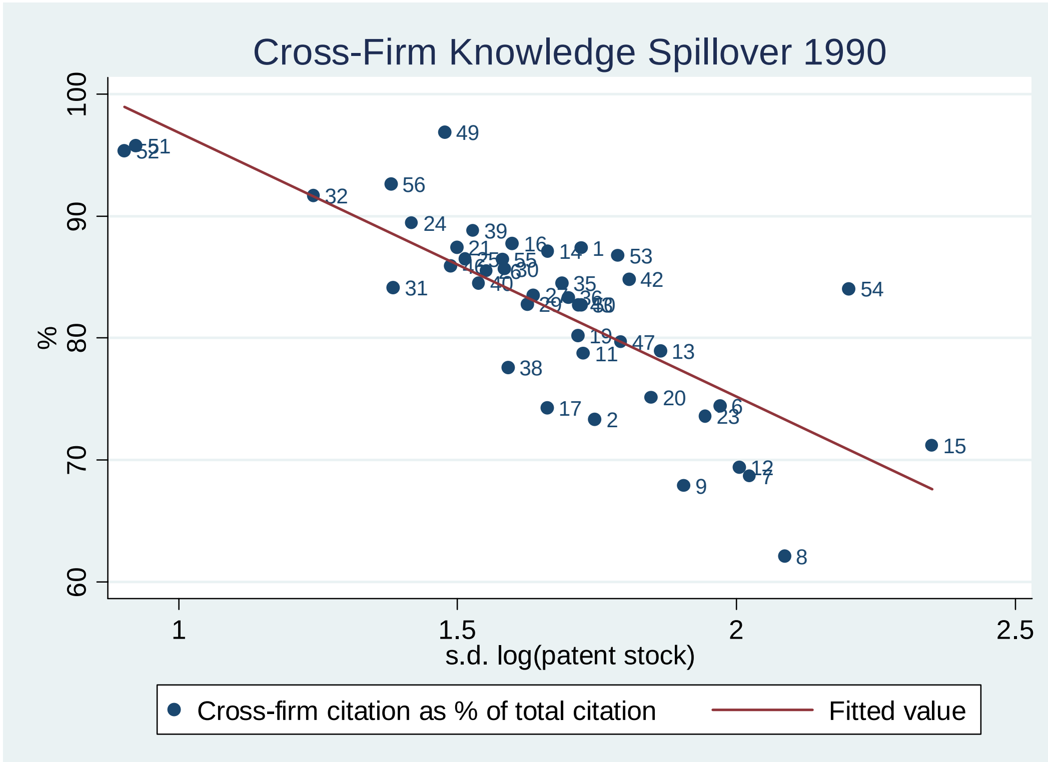

Figure 9

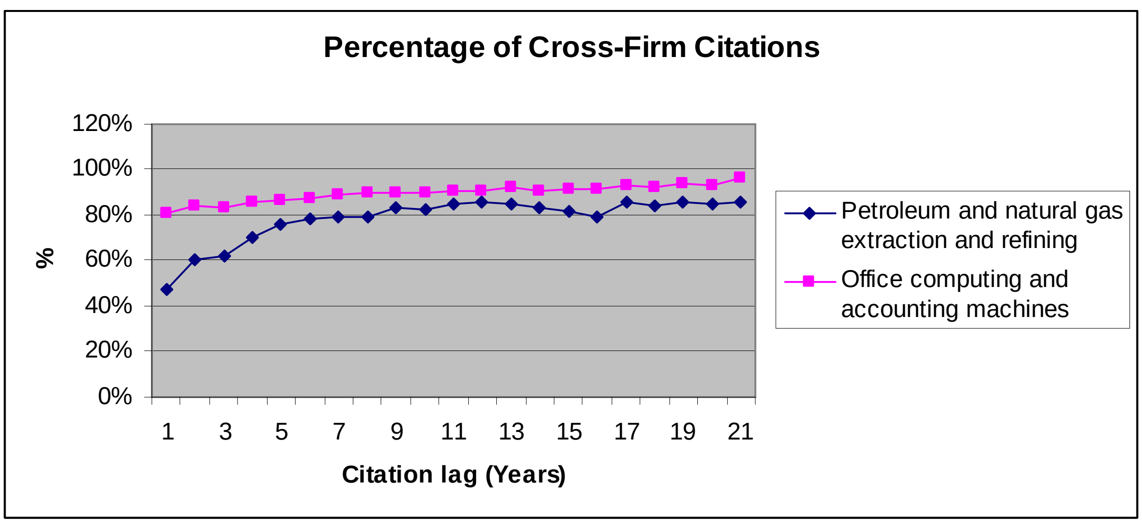

Figure 10

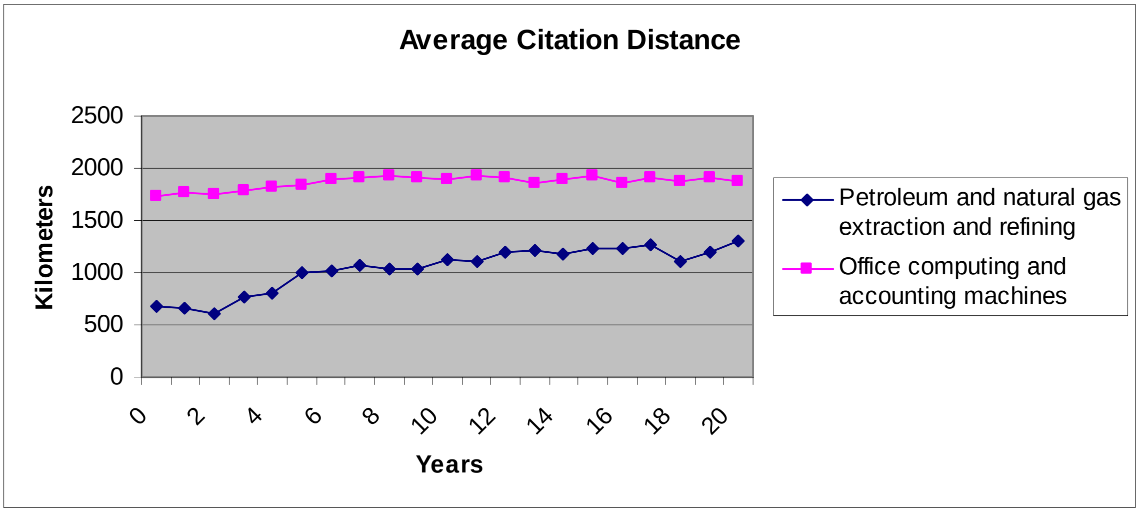

Figure 11

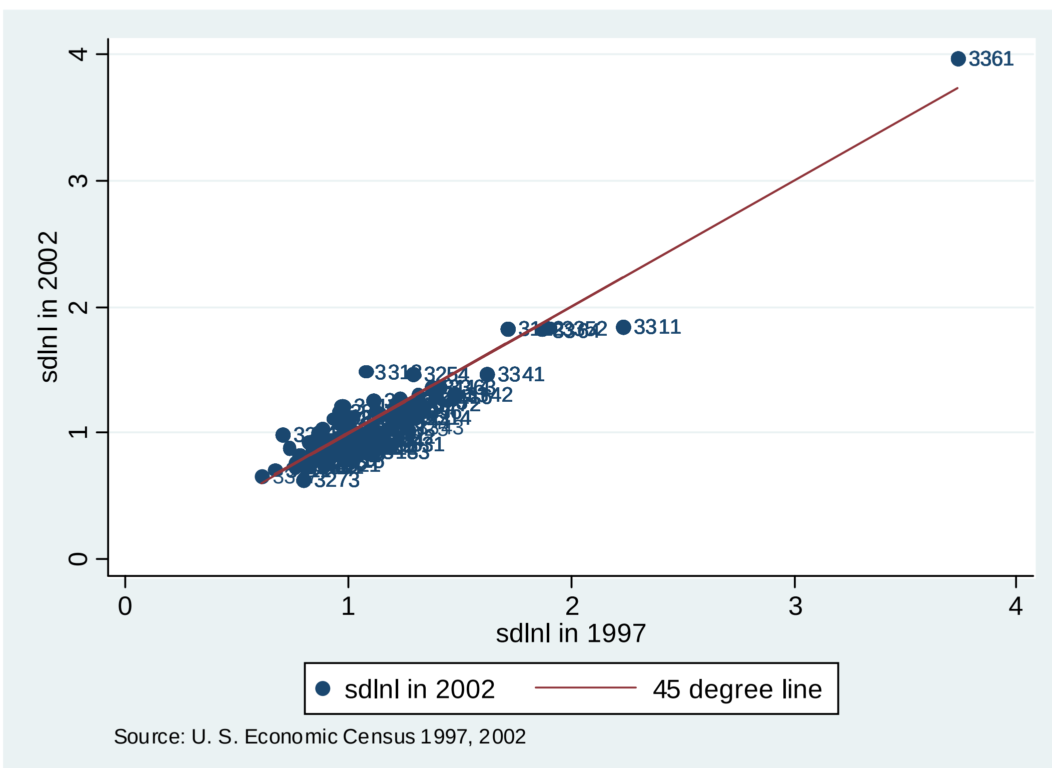

Figure 12: Sectoral Firm Size Heterogeneity in U.S. 1997 and 2002

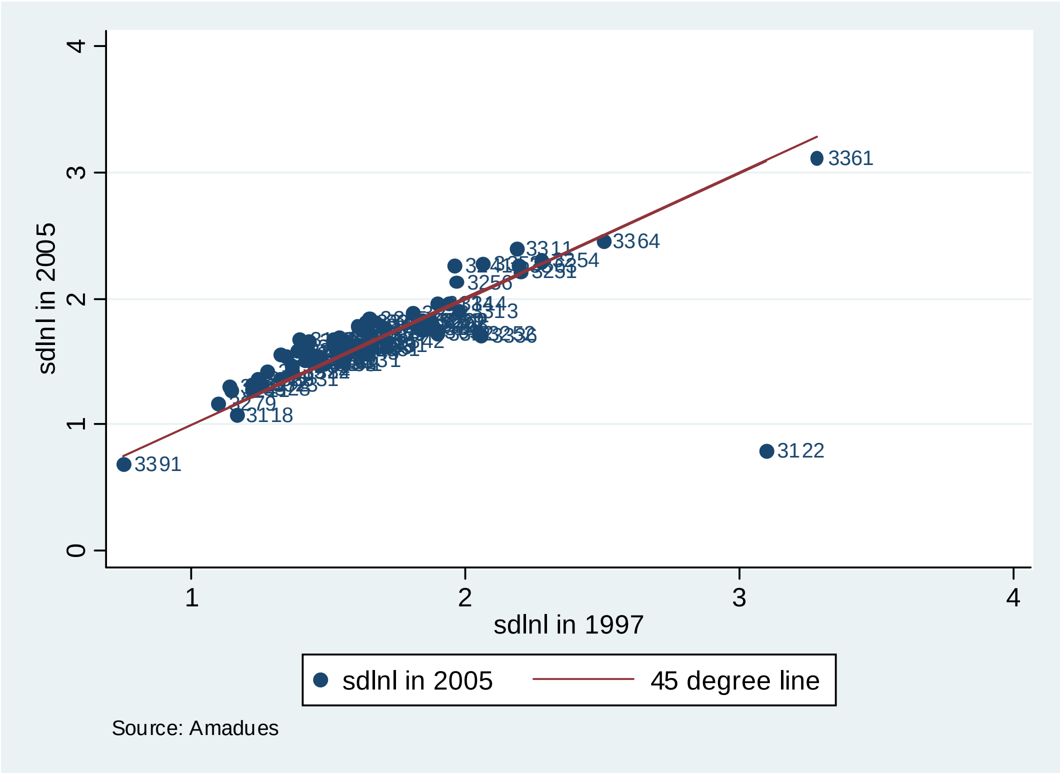

Figure 13: Sectoral Firm Size Heterogeneity in France 1997 and 2005

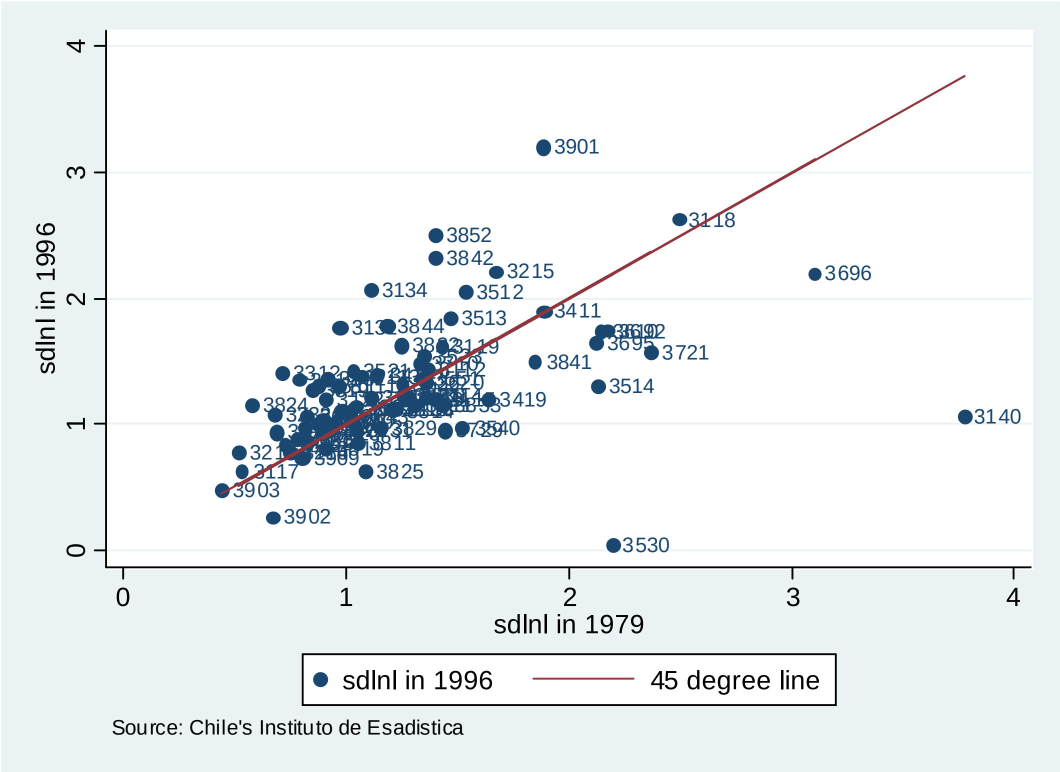

Figure 14: Sectoral Firm Size Heterogeneity in Chile 1979 and 1996

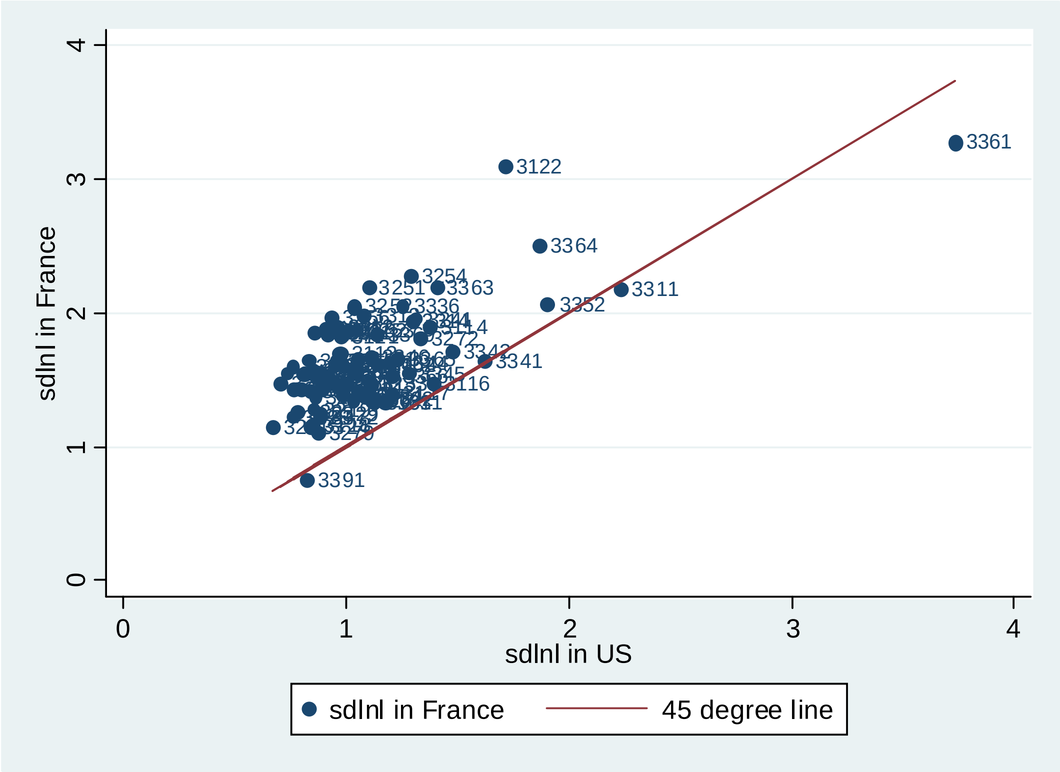

Figure 15: Sectoral Firm Size Heterogeneity in U.S. and France 1997

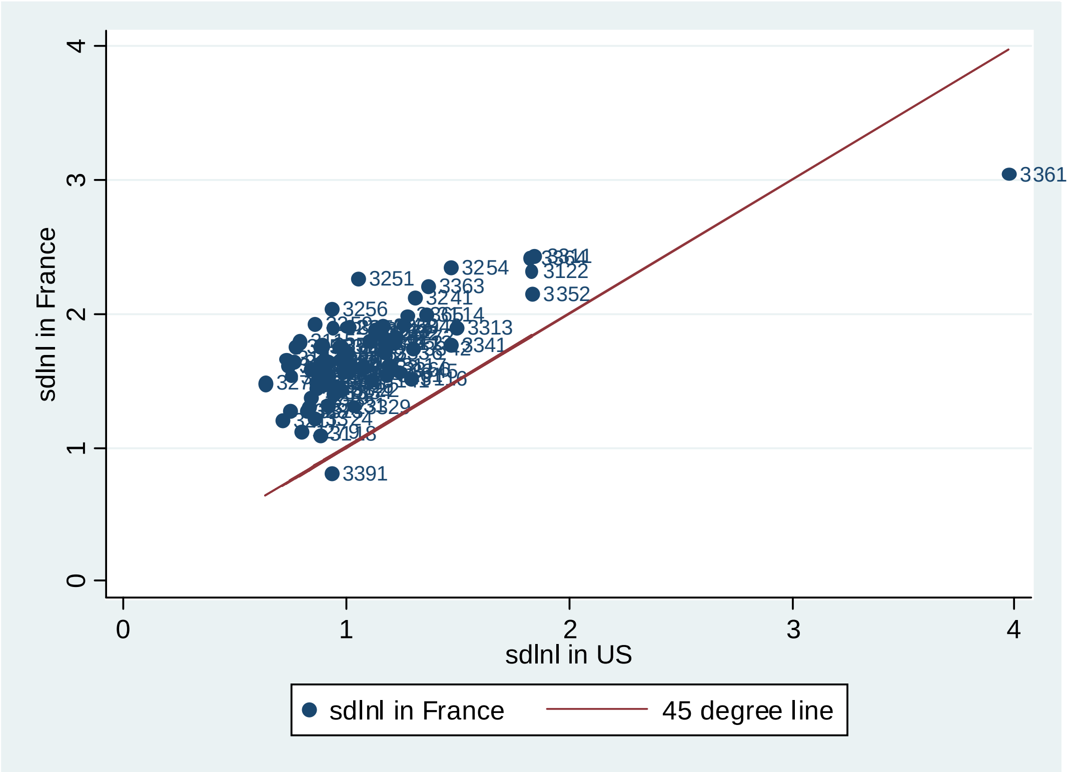

Figure 16: Sectoral Firm Size Heterogeneity in U.S. and France 2002

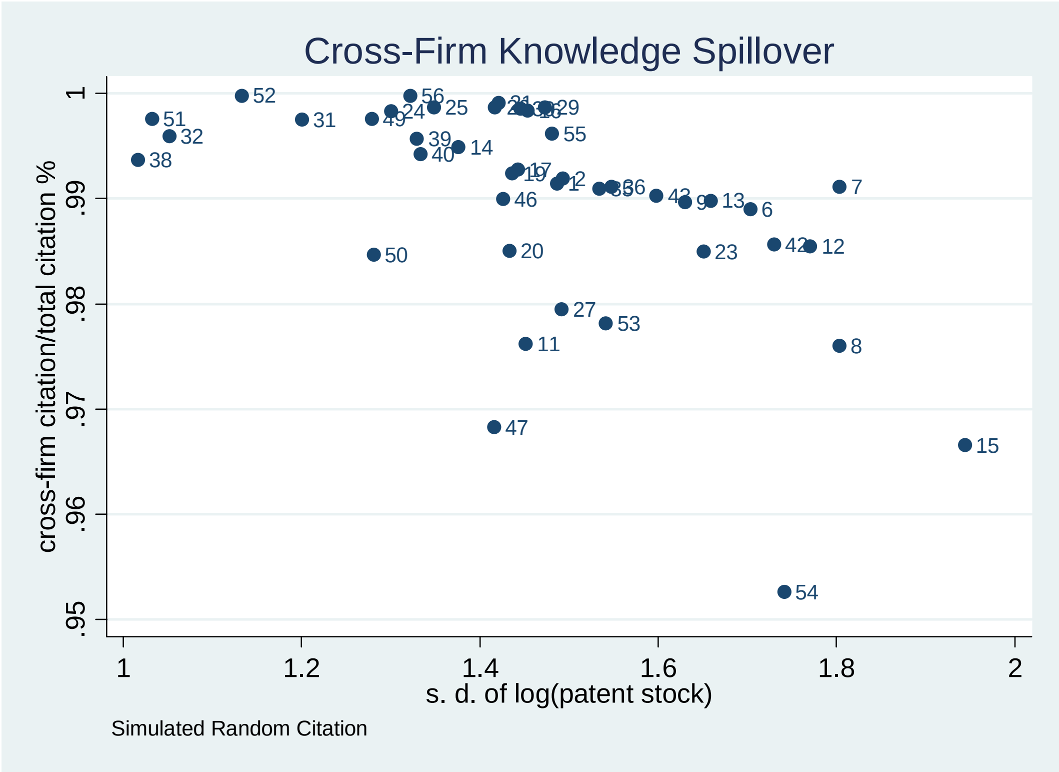

Figure 17

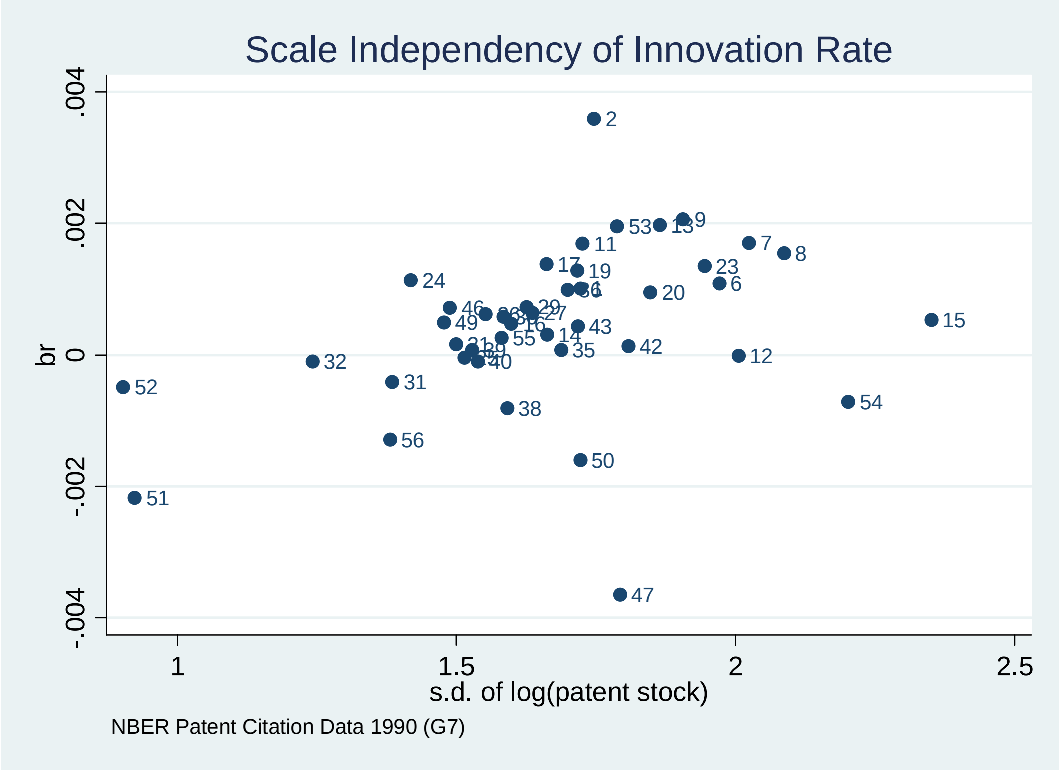

Figure 18

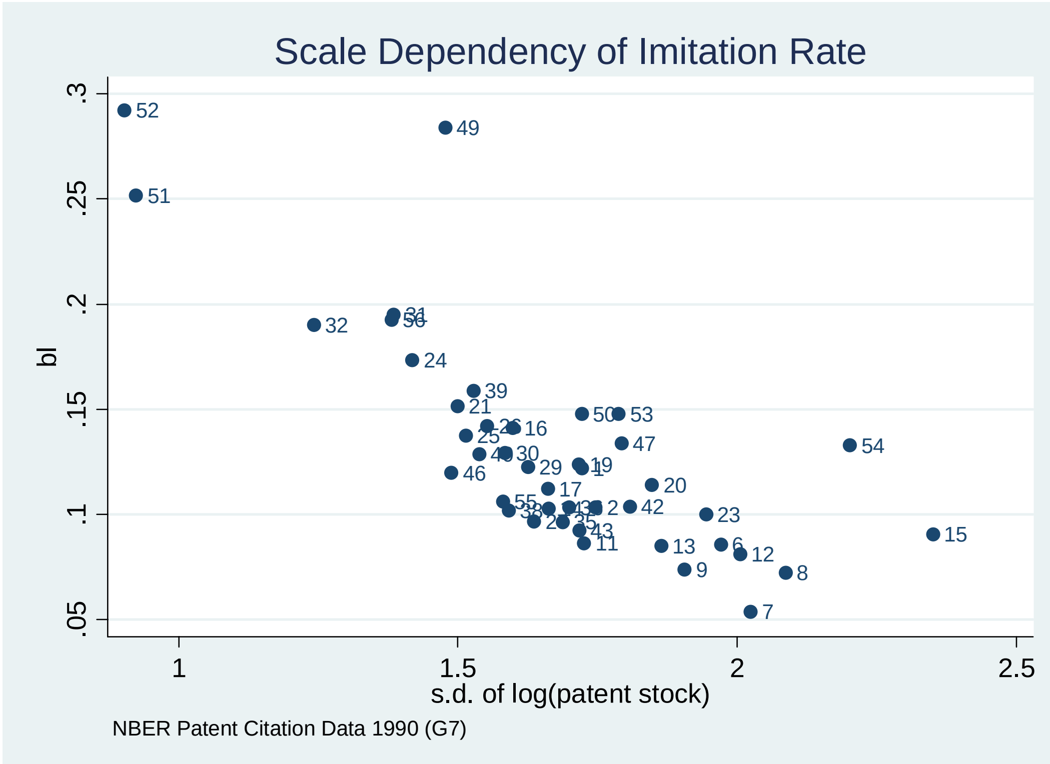

Figure 19

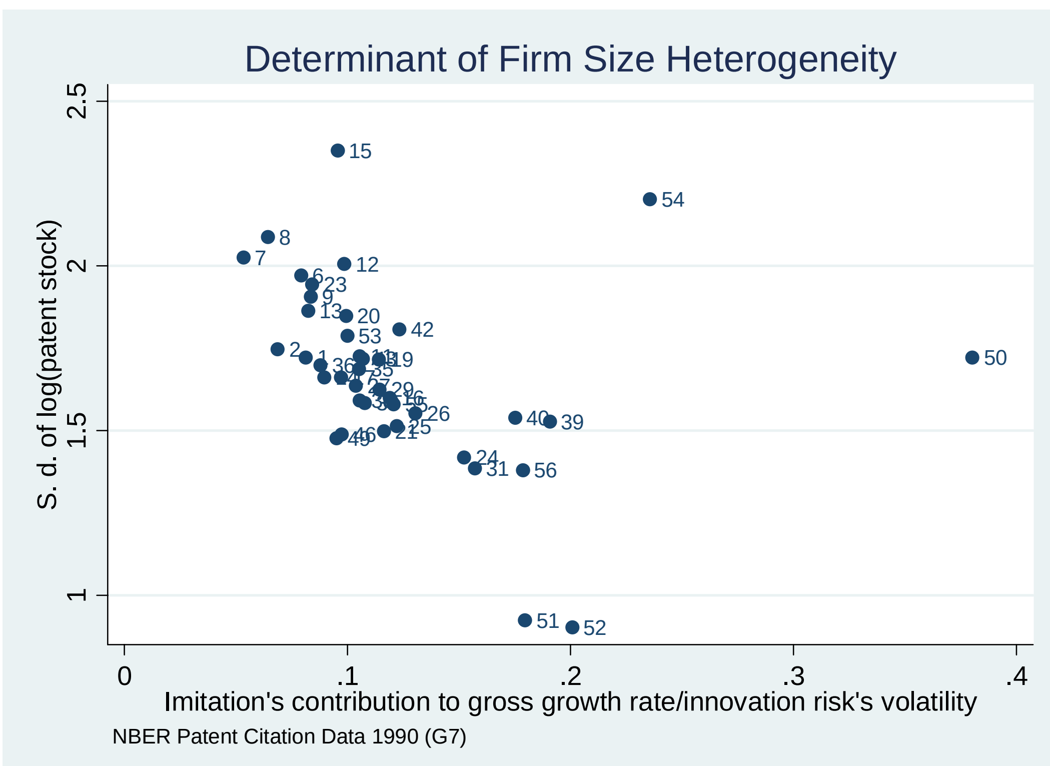

Figure 20

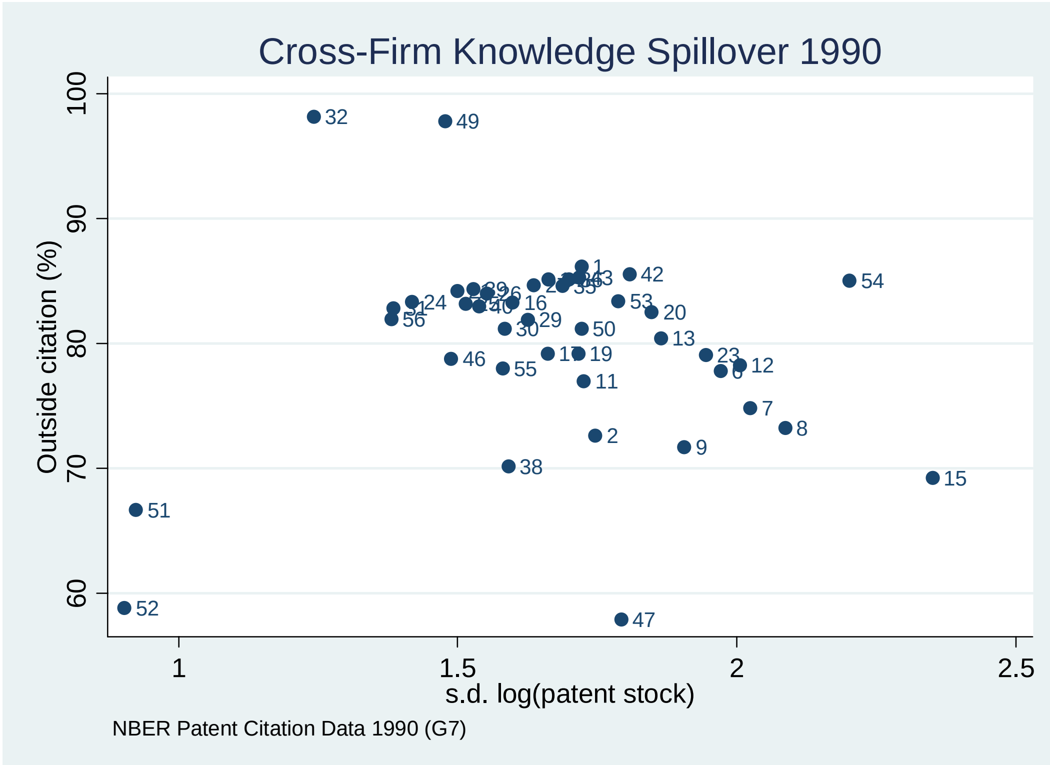
Cross-Firm Knowledge Spillover 1990
Outside citation (%)
s.d. log(patent stock)
NBER Patent Citation Data 1990 (G7)

Figure 21

OLS Regressions with U.S. Citations Dependent variable: citation lag				
Independent variable	1	2	3	**4**
S. d. of log(ps)	.776** (.372)	2.934** (1.236)	1.262** (.563)	2.443*** (.455)
Log(dist)	.137*** (.017)	.122*** (.013)	.142*** (.008)	.030*** (.010)
Log(citing_ps)	-.147*** (.008)	-.110*** (.021)	-.099*** (.004)	.555*** (.063)
Log(cited_ps)	-.152*** (.008)	-.080*** (.014)	-.113*** (.005)	-2.621*** (.146)
Log(sic_ps)	-.355*** (.014)	-3.470*** (.429)	-3.215*** (.125)	1.208*** (.128)
Fixed effects	State pair	Sector	State pair by sector	Firm pair
No. of observations	1132505	1132505	1132505	1132505
No. of groups	2626	42	47375	719298
R square	.095	.093	.173	.855

Robust standard errors clustered by sector are reported in the brackets. Year dummies are included.
Log(dist) is the log scale great circle distance between the citing patent and the cited patent.
Log(citing_ps) (Log(cited_ps)) is the log scale patent stock of the citing (cited) firm.
Log(sic_ps) is the log scale patent stock of the sector.
S. d. of log(ps) is the standard deviation of log scale patent stock for all firms in the sector.

Table 1a

Summary of Variables					
	Obs.	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Citation lag	1132640	7.289	5.231	0	94
Log(dist)	1132640	6.711	1.784	0	9.460
Log(citing_ps)	1132640	2.539	2.574	0	10.490
Log(cited_ps)	1132640	2.769	2.618	0	10.280
Log(sic_ps)	1132640	10.611	1.052	4.111	12.382
S. d. of log(ps)	1132640	1.667	0.217	0.786	8.495

Table 1b

Correlation between Variables						
	Citation lag	**Log(dist)**	**Log(citing_ps)**	**Log(cited_ps)**	**Log(sic_ps)**	**S. d. of log(ps)**
Citation lag	1					
Log(dist)	0.0304	1				
Log(citing_ps)	-0.1211	0.0133	1			
Log(cited_ps)	-0.1106	-0.0137	0.371	1		
Log(sic_ps)	-0.0605	0.055	0.2123	0.2123	1	
S. d. of log(ps)	0.0511	-0.021	0.1906	0.2139	0.0942	1

Table 1c

TABLE 2
Average Firm Characteristic by Firm Size (2003Q4)

Value Bins (in $)	UPCs	Brands	Modules	Product Groups	Share
1- 100,000	3	1	1	1	0.00
100,000 - 1,000,000	10	2	3	2	0.02
1,000,000 - 5,000,000	33	5	6	3	0.05
5,000,000 - 10,000,000	69	7	10	4	0.04
10,000,000 - 50,000,000	139	11	20	7	0.17
50,000,000 - 100,000,000	386	22	50	14	0.10
100,000,000 - 500,000,000	713	37	71	18	0.33
500,000,000 - 1,000,000,000	1191	68	110	28	0.12
> 1,000,000,000	3431	182	246	54	0.16

Note: The share is based on total value of UPCs for firms within a bin compared to total value in 2003Q4
Ignore entry and exit of firms, only consider firms that exists in both periods t and t-1

Table 2: From Broda and Weinstein (2007)

Number of Patents	Average Number of Patent Categories
1 - 10	1.34
11 - 100	3.89
101 - 1000	8.93
1001 - 10000	15.17
- 10000	25.57
Source: NBER Patent Citation Data 1999	

Table 3

The Citing Sector	The Cited Sector						
	%	1	2	6	7	8	9
	1	**80.34%**	0.00%	0.00%	6.74%	0.00%	8.15%
	2	0.33%	**38.59%**	0.33%	0.66%	8.70%	0.00%
	6	0.11%	0.42%	**60.30%**	10.72%	1.27%	0.42%
	7	0.46%	0.41%	5.16%	**58.46%**	4.52%	14.06%
	8	0.00%	1.44%	1.32%	7.38%	**66.33%**	0.06%
	9	1.09%	0.30%	0.30%	14.68%	0.24%	**67.73%**

Table 4a Cross-Sector Citations

The Citing Sector	The Cited Sector						
	Adjusted%	1	2	6	7	8	9
	1	**121.1008**	0	0	0.696048	0	**2.422689**
	2	0.495044	**36.2602**	0.131709	0.067814	**2.784302**	0
	6	0.160022	0.399015	**24.18251**	**1.107003**	0.407557	0.126287
	7	0.694979	0.389908	**2.070915**	**6.035842**	**1.445521**	**4.182057**
	8	0	**1.353679**	0.529604	0.762269	**21.21998**	0.017852
	9	**1.639478**	0.283891	0.121165	**1.515955**	0.077325	**20.14454**

Table 4b Cross-Sector Citations Adjusted by the Cited Sector's Weight

1	Food and kindred products
2	Textile mill products
6	Industrial inorganic chemistry
7	Industrial organic chemistry
8	Plastics materials and synthetic resins
9	Agricultural chemicals

Table 4c Sector Name

Year	corr(sdl,sdy)	corr(sdl,sds)	corr(sdl,sdva)	corr(sds,sdy)	corr(sds,sdva)	corr(sdy,sdva)
1997	0.964	0.961	0.970	0.998	0.964	0.963
1998	0.954	0.956	0.961	0.998	0.940	0.937
1999	0.951	0.954	0.967	0.997	0.918	0.912
2000	0.924	0.926	0.926	0.998	0.944	0.942
2001	0.959	0.959	0.952	0.998	0.933	0.929
2002	0.956	0.962	0.915	0.997	0.911	0.906
2003	0.924	0.922	0.945	0.997	0.931	0.928
2004	0.911	0.926	0.937	0.915	0.943	0.888
2005	0.920	0.933	0.937	0.996	0.920	0.907

Table 5: Correlation between Firm Size Heterogeneity Measures by Different Firm Size Proxies in French Data

Year	corr(sdl,sdy)	corr(sdl,sds)	corr(sdl,sdva)	corr(sds,sdy)	corr(sds,sdva)	corr(sdy,sdva)
1979	0.8503	0.8268	0.8455	0.9659	0.9422	0.9614
1980	0.8486	0.7744	0.8139	0.9076	0.8989	0.9490
1981	0.8543	0.7933	0.7768	0.9546	0.8919	0.9131
1982	0.8259	0.7208	0.7436	0.9353	0.9519	0.9406
1983	0.7007	0.6433	0.6938	0.9571	0.9142	0.9473
1984	0.7327	0.7351	0.7133	0.9684	0.9446	0.9770
1985	0.7712	0.7341	0.7534	0.9628	0.9418	0.9617
1986	N. A.	N. A.	N. A.	N. A.	N. A.	N. A.
1987	0.7853	0.7249	0.7454	0.9542	0.8965	0.9185
1988	0.8057	0.7243	0.7312	0.9293	0.8865	0.9216
1989	0.8407	0.7965	0.8173	0.9553	0.8838	0.9338
1990	0.8033	0.7532	0.7790	0.9675	0.8875	0.9328
1991	0.6233	0.6186	0.5515	0.9602	0.8958	0.9114
1992	N. A.	N. A.	N. A.	N. A.	N. A.	N. A.
1993	N. A.	N. A.	N. A.	N. A.	N. A.	N. A.
1994	0.6678	0.5916	0.6190	0.8591	0.8579	0.9233
1995	0.7383	0.6122	0.3697	0.8491	0.7765	0.7468
1996	0.8254	0.7687	0.7150	0.9328	0.8958	0.9103

Table 6: Correlation between Firm Size Heterogeneity Measures by Different Firm Size Proxies in Chilean Data

OLS Regressions with Random Citations Dependent variable: citation lag (Random Citation Data)			
Independent variable	1	2	3
S. d. of log(ps)	.424*** (.059)	.160** (.078)	.168*** (.066)
Log(dist)	.065 (.042)	-.051 (.026)	0.039 (0.026)
Log(citing_ps)	-.040*** (.005)	-.002 (.002)	-.001(.002)
Log(cited_ps)	-.170*** (.010)	-.108 (.036)	-.160*** (.007)
Log(sic_ps)	-.175*** (.014)	-3.603*** (.179)	-3.341*** (.096)
Year fixed effect	Yes	Yes	Yes
State pair fixed effects	Yes	No	No
Industry fixed effects	No	Yes	No
State pair - industry fixed effects	No	No	Yes
No. of observations	2120904	2120904	2120904
No. of groups	2626	42	47375

Robust standard errors clustered by sector are reported in the brackets. Year dummies are included.
Log(dist) is the log scale great circle distance between the citing patent and the cited patent.
Log(citing_ps) (Log(cited_ps)) is the log scale patent stock of the citing (cited) firm.
Log(sic_ps) is the log scale patent stock of the sector.
S. d. of log(ps) is the standard deviation of log scale patent stock for all firms in the sector.

Table 7a

Summary of Variables (Random Citation Data)					
	Obs.	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Citation lag	2120904	8.332	6.385	0	94
Log(dist)	2120904	7.043	1.189	0	9.530
Log(citing_ps)	2120904	3.473	2.695	0	9.649
Log(cited_ps)	2120904	3.134	2.667	0	9.563
Log(sic_ps)	2120904	10.645	1.055	4.111	12.382
S. d. of log(ps)	2120904	1.700	.243	.787	8.495

Table 7b

Correlation between Variables (Random Citation Data)						
	Citation lag	**Log(dist)**	**Log(citing_ps)**	**Log(cited_ps)**	**Log(sic_ps)**	**S. d. of log(ps)**
Citation lag	1					
Log(dist)	-.009	1				
Log(citing_ps)	-.031	-.047	1			
Log(cited_ps)	-.062	-.040	.172	1		
Log(sic_ps)	.020	.078	.211	.219	1	
S. d. of log(ps)	.054	-.042	.190	.199	.048	1

Table 7c

OLS Regressions with G7 Citations Dependent variable: citation lag				
Independent variable	1	2	3	4
S. d. of log(ps)	.576 (.353)	.492 (1.164)	1.99** (.774)	1.917*** (.440)
Log(dist)	.078*** (.020)	.122*** (.050)	.104*** (.022)	.044*** (.007)
Log(citing_ps)	-.149*** (.014)	-.134*** (.008)	-.112*** (.008)	.562*** (.050)
Log(cited_ps)	-.183*** (.044)	-.179*** (.059)	-.149*** (.051)	-2.464*** (.164)
Log(sic_ps)	-.335*** (.089)	-2.787*** (.365)	-2.939*** (.373)	1.215*** (.104)
Fixed effects	State pair	Sector	State pair-sector	Firm pair
No. of observations	2158761	2158761	2158761	2158761
No. of groups	49	42	1884	1238745
R square	.089	.098	.111	.805

Robust standard errors clustered by sector are reported in the brackets.
Log(dist) is the log scale great circle distance between the citing patent and the cited patent.
Log(citing_ps) (Log(cited_ps)) is the log scale patent stock of the citing (cited) firm.
Log(sic_ps) is the log scale patent stock of the sector.
S. d. of log(ps) is the standard deviation of log scale patent stock for all firms in the sector.

Table 8a

Summary of Variables (G7 Citation Data)					
	Obs.	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Citation lag	2158761	6.866	5.010	0	95
Log(dist)	2158761	7.310	2.036	0	9.760
Log(citing_ps)	2158761	3.321	2.778	0	10.491
Log(cited_ps)	2158761	3.451	2.689014	0	10.280
Log(sic_ps)	2158761	10.651	1.033	4.111	12.382
S. d. of log(ps)	2158761	1.687	.231	.786	8.494

Table 8b

Correlation between Variables (G7 Citation Data)						
	Citation lag	**Log(dist)**	**Log(citing_ps)**	**Log(cited_ps)**	**Log(sic_ps)**	**S. d. of log(ps)**
Citation lag	1					
Log(dist)	0.038	1				
Log(citing_ps)	-0.155	0.035	1			
Log(cited_ps)	-0.162	0.028	0.423	1		
Log(sic_ps)	-0.058	0.025	0.273	0.278	1	
S. d. of log(ps)	0.026	0.024	0.157	0.181	0.069	1

Table 8c