Unleashing the Dragon: The Case for Patent Reform in China

Lee G. Branstetter* Douglas Hanley† Huiyan Zhang‡

February 8, 2023

Abstract

This paper explores key features of China's patent system, embedding them in a growth model that allows us to explore their consequences for long-run innovation, economic growth, and welfare. We first show that China's system is characterized by narrow patent protection, reflecting three key features: a strong bias towards incremental innovation, weak patent enforcement, and a decline in the quality of patent examination. Based on these stylized facts, we build a growth-theoretic model to show how narrow patent protection distorts innovation incentives. By skewing R&D efforts toward "incremental" innovation, China's patent system slows technological progress and lowers national economic growth. These results suggest that strengthening patent infringement damage awards and raising patent application fees could significantly accelerate China's economic and technological progress.

Keywords— Patents, China, Endogenous Growth, Innovation

*Carnegie Mellon University
†University of Pittsburgh
‡Carnegie Mellon University; Católica Lisbon School of Business and Economics

1 Introduction

Although China has become the world's largest generator of patented inventions, the quality of its intellectual property system remains the subject of considerable controversy inside and outside of China. A recent wave of (mostly empirical) research has documented specific shortcomings of Chinese patent policy (Chen et al., 2021; Dang and Motohashi, 2015; Li, 2012; Wei et al., 2021; Sun et al., 2021), but more work is needed to analyze the consequences of these failings for innovation within China. This paper intends to address these gaps by documenting key features of China's patent system and embedding them in a theoretical model that facilitates exploration of the consequences for innovation, economic growth and social welfare.

We start with a description of China's patent system that documents some important facts and points to three key conclusions. First, Chinese patented inventions appear to skew heavily in the direction of incremental rather than more substantive inventions. We argue that this distribution reflects the inadequate protection offered to broad, substantive, or "radical" inventions. Second, we show that the legal sanctions that punish patent infringement in China are extremely weak compared to other major jurisdictions. This leads to high rates of infringement, especially for more socially valuable, significant inventions. Low damages for infringement reinforce the incremental nature of Chinese innovation. Third, we show evidence suggesting that China's recent tidal wave of patent applications, combined with a political directive to limit patent examination delays, has weakened the ability of the Chinese patent system to apply a consistent standard of quality and novelty across inventions. The surge of applications has itself been driven, in part, by policy-driven subsidies that have reduced the cost of patenting to very low levels. This makes it all but impossible for China's hardworking patent examiners to distinguish adequately between incremental and more substantive inventions or to draw intelligent boundaries between related patented ideas. Like low damage awards, inadequate examination reinforces the incremental nature of Chinese invention by limiting the degree to which truly innovative breakthroughs are distinguished from minor modifications of prior inventions.

Why does this matter? Because national inventive progress is constrained when big leaps in technological progress are insufficiently protected. We show this with a growth-theoretic model in which Chinese patent policy determines the nature and extent of firm innovative efforts. Any Chinese firm that takes on the higher risk and cost of substantive invention and actually succeeds in producing a much higher quality product, process, or service is quickly imitated. Low infringement damage awards are insufficient to deter infringement of high quality inventions. Infringement lowers the net benefit of "radical" innovation and leads firms to focus on incremental innovations that are less costly and risky, but yield less of an inventive step. In equilibrium, this leads to slower technological progress for China, lowering national economic welfare. We calibrate the theoretical model using real-world data, obtaining key parameters of the model, such as the research aptitude distribution, the distribution of innovation step size and the innovation rate.

We then consider the implications of our model for future Chinese patent policy. We argue that the effective breadth or scope of protection for substantive or radical inventions must be increased. How can this be achieved? First, the distribution of patent damage awards must shift significantly to the right. Second, to truly become the innovation superpower it aspires to be, China needs fewer patents but better ones. Higher application fees could reduce the flow of duplicative, low-quality inventions and simultaneously provide Chinese patent examiners with the resources needed to engage in a more careful review of these applications. Movement in these directions will benefit China, expand technological possibilities for the entire world, and do much to address the trade frictions that beset China's economic and political interactions with its most important trading partners. China's growing share of the world's innovation resources makes these reforms an urgent priority for the entire human race.

This paper contributes to several lines of research. First, our study is closely related to a recent wave of empirical research on Chinese patent policy (Chen et al., 2021; Dang and Motohashi, 2015; Li, 2012; Wei et al., 2021; Sun et al., 2021; Hu and Jefferson, 2009; Hu, 2014). Building on the shortcomings of the Chinese patent system documented by these studies, we take the crucial step of embedding the policy problems and distortions in a growth theoretical model and investigating the implication of these problems on China's innovation system and the economic growth. Our theoretical model and calibration of the model to data allow us to quantify how and to what extent China's patent system lowers the nation's welfare by limiting the kinds of innovations Chinese firms choose to pursue. Finally, our growth-theoretic model points us to specific reforms of the patent system that could substantially improve Chinese and, by implication global economic well-being.

This paper also contributes to the extensive theoretical and empirical literature exploring the role of patent protection in incentivizing innovation (Qian, 2007; Hall and Ziedonis, 2001; Branstetter and Sakakibara, 2002; Moser, 2005; Lerner, 2009). Prior research has explored optimal patent length and breadth (Nordhaus, 1969; Klemperer, 1990; Gilbert and Shapiro, 1990). Our research has been particularly influenced by the work of O'Donoghue et al. (1998), who note the distinction between lagging and leading patent breadth and consider the optimal design of patents in terms of both length and breadth when facing a stochastic innovation process. Hopenhayn and Mitchell (2001) and Cornelli and Schankerman (1999) build on this by also considering the possibility of having a menu of different patent types. Other relevant work in this research stream has explored the role of patents in exhibiting innovators' capabilities (Agarwal et al., 2009; Ganco et al., 2015) and improving welfare gains by facilitating the market for technology (Arora et al., 2004; Arora and Gambardella, 1994). Prior research has also considered the challenges faced by countries with relatively weak intellectual property rights (IPR) regimes, where weak enforcement hampers technological advance (Lamin and Ramos, 2016; Belderbos et al., 2021; Branstetter et al., 2006). As a result, innovating firms may need alternative innovation strategies to protect their inventions (Zhao, 2006; Beukel and Zhao, 2018; Paik and Zhu, 2016). We demonstrate that economies with weak IPR regimes could substantially benefit from improvement in patent protection.

Finally, our paper builds on the insights and mechanism developed in a number of different areas of the applied theoretical literature on economic growth. The model employed is one of fully endogenous growth, and so builds on the seminal contributions of both Grossman and Helpman (1991) and Aghion and Howitt (1992), particularly in the importance of creative destruction to the growth process. In addition, a number of more recent papers have studied heterogeneity in the innovation process along a variety of dimensions: Akcigit and Kerr (2018) look at the dichotomy between internal innovation to improve a firm's own products and external innovation to capture new product lines; Acemoglu et al. (2018) consider a setting in which firms vary in their intrinsic aptitude for generating innovations; Akcigit et al. (2021) study a setting in which firms choose between two types of innovation technologies, applied and basic. We focus on the distinction between incremental and radical modes of innovation, an area in which there is already some existing research. In particular, Acemoglu et al. (2020) study a setting in which firms can choose between incremental and radical innovation and consider the properties of managers (primarily age) that make them more or less likely to engage in radical innovation. We differ from this work in that we focus primarily on how the nature and design of patent policy influence firms' propensity to undertake radical innovation.

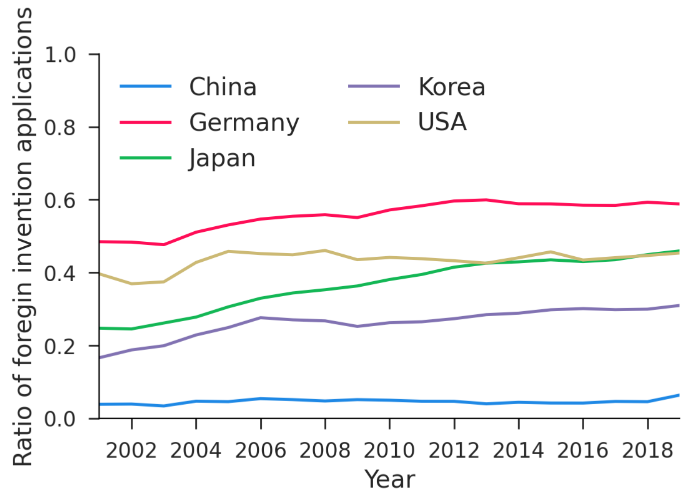
2 Key Features of China's Patent System

2.1 A strong bias toward incremental innovation

China did not establish a modern patent system until 1985. It has witnessed rapid growth in patent applications since then. However, a variety of indicators suggest that the quality of patents filed in China is far less impressive than their quantity. The growth in invention patents—the type of patents which undergo a formal examination process, analogous to US utility patents—is dwarfed by growth in utility model patents—a form of intellectual property protection for more incremental inventions whose grants are not contingent on any examination of the novelty of the invention. This suggests that China's vaunted patent explosion is mostly driven by incremental inventions.

If we set utility models aside, we still see impressive growth in the number of invention patents applied for and granted in China. However, there is strong evidence that even these Chinese inventions, which go through a formal evaluation process that supposedly confirms their novelty, are of far lower quality than their foreign counterparts. Figure 1 shows the percentages of domestic patents filed by inventors in five major countries which

Figure 1: Fraction of Invention Patents Filed Abroad by Country



Source: WIPO

forego foreign patent protection. The percentages in Germany, the US, Japan, and South Korea are 60%, 40%, 20-40%, and around 20%. In striking contrast, Chinese inventors only obtain foreign patent protection for 5% of their domestic inventions, and this fraction has not budged in recent years even as Chinese patenting has exploded and measured R&D expenditures have grown sharply. This might be reasonable if Chinese firms had no economic interests abroad, but China is the world's largest exporter of goods. Despite this deep integration with the global economy, Chinese inventors clearly view about 95% of their invention patents as being not worth patenting in a single nation outside of China.

2.2 Weak patent enforcement

The second major point we wish to make is that enforcement of patent rights in China remains, despite decades of patent reform, very weak by international standards. As many patent lawyers point out, strong patent systems confer on patent owners not a shield but a sword—the sword is the ability to sue infringers for damages sufficiently large to deter patent infringement, or credibly threaten to do so. In China, infringement damages remain far too low to have deterrent value. Although statutory infringement compensation was enhanced, in principle, to the range of ten thousand to one million RMB

Table 1: Summary Statistics for Patent Infringement Lawsuits

	Invention	Utility	Design	Total
Average damage awards (2021 USD)	68,000	16,838	7,606	17,400
Average proceeding time (months)	8.6	6.2	5.1	5.9
Compensation received / claimed	38%	28%	29%	33%
Win rate	63%	71%	83%	77%

Source: CIELA, https://www.ciela.cn/en/analysis/patents.

(USD 1,553 to 155,300) by the patent reform implemented in 2008 [1], in practice, the average compensation received by invention patent owners who seek infringement remedies is only around 440,000 RMB (USD 68,000).[2] This could be less than the amount of money it would take to buy a single high-end luxury car in China in which to transport the executives of a single patent-infringing firm. This small amount accounts for only 38 percent of the infringement compensation claimed by the patent owner, on average, and is less than half of the statutory damage awards stipulated by the law, as shown in Table 1. The average legal compensations for infringement of utility model and design patents are even lower—108,388 RMB (USD 16,838) and 48,965 RMB (USD 7,606) respectively. Even when courts rule in favor of the incumbent patent holder, these decisions are imperfectly enforced, and Chinese inventors often fail to collect what the courts have awarded them.

A comparison with other major patent systems demonstrates that China's system is characterized by extremely low damage awards. Table 2 shows a comparison of China, Japan and the US in terms of major features of patent litigation practice. At every point of the distribution, damage awards in China are only a small fraction of those in the U.S. For instance, the average and median of damage awards received by Chinese patent owners are merely 0.3 percent and 0.9 percent of those received by US patent owners, and at the 75th percentile and the maximum, damage awards are still well below 1 percent of their U.S.

[1] China enacted its newest version of patent law in 2021, which stipulates the statutory damage awards to be from USD 4660 to USD 776500. Though our discussion is mostly restricted to the period before the newest round of patent reform, the new damage awards range is still low by international standards.

[2] This data is collected from CIELA, a private firm providing patent litigation data. See https://www.ciela.cn/en/.

Table 2: Comparison of Patent Litigation Statistics in China, Japan, and the U.S.

	China	Japan	US
Damage awards (2010 US$)			
min	1,291	44	1,873
25th percentile	15,838	27,201	500,000
mean	67,825	1,282,778	18,879,257
median	31,515	170,429	3,196,139
75th percentile	65,425	806,029	13,608,450
max	1,500,150	22,068,481	575,283,461
Avg proceeding time (months)	6-18	12-15	18-42
Avg lawsuit costs (2010 US $)	20k-150k	300k-500k	1m-6m
Avg damage awards / avg lawsuit costs	3.4-0.45	4.3-2.6	18.9-3.1
Win rate	77%	21.80%	53%-76%

Note: *Chinese win rate is from CIELA data. Japanese win rate is recalculated from Yamaguchi (2010); data window is from 2010 to 2014. US win rate is from Barry et al. (2017); the data collection window is from 2011 to 2016. Distribution of damage awards is from Hu et al. (2020). The rest of the figures are from a WIPO report (Helmers et al., 2018).*

levels. The patent system adopted by China in the 1980s was based on Japan's system at the time, and even now Japan's patent litigation system is characterized by a relatively low level of damage awards (Helmers et al., 2018). However, China's distribution of damage awards lies far to the left of Japan's. To put it bluntly, damage awards in China are far too low to deter rampant infringement (Economist, 2020).

Defenders of China's patent enforcement system often point to other dimensions in which it fares better in international comparison. The average time from initiation to initial judgment for first-instance patent infringement cases in China is merely half of that in Japan, and one-third of the average in the US. The average monetary cost (court fees and lawyer compensation) for a standard infringement proceeding is also much lower in China (Table 2). However, when taking into account the average damage awards divided by the average lawsuit costs (the last row of Table 2), China's ratio ranges from 3.4 to 0.45, while these ratios for Japan and the US are all above one, suggesting that the expected returns to Chinese patent owners from pursuing a patent lawsuit could be negative, even when they "win." Defenders of the Chinese system often point out that the "win rate" in China—the probability that patent owners successfully defend their patents against infringers in court—is comparable to that in the US and significantly higher than the rate observed in Japan. However, this matters little if what is "won" is a paltry judgment that is little more than a minor inconvenience for a determined infringer. Finally, one might also argue that an injunction to infringement is awarded in almost all winning cases. The actual effects are reduced by another chronic problem: judicial rulings of damage awards and injunctions are often not fully enforced. For some patent holders, winning an infringement lawsuit might only be the start of another frustrating process—ensuring that damage awards get paid in a reasonable amount of time and, more importantly, infringement stops. In any case, these efforts to paint Chinese patent litigation trends in a positive light are powerfully undermined by a more comprehensive view of the data. The magnitude of China's infringement rate, even at its peak of around 0.13%, is much lower than the corresponding level in other countries. (e.g. the litigation rate in the US is reported to be around 4% (Bessen and Meurer, 2005)). The low and rapidly declining infringement rate in China is understandable if infringement damages are too low to deter infringement.

2.3 Deterioration of patent examination

The next major characteristic of China's patent system on which we wish to focus is the deterioration of the examination process. China is perhaps unprecedented among major patent jurisdictions in submitting the vast majority of the patents it issues—utility models—to no examination. The other major patent jurisdictions that grant utility models, such as Japan and Germany, do so sparingly. In striking contrast, China now grants more than three times as many utility models as invention patents, and Chinese courts do not treat these two categories of patents that differently. Utility models expire sooner than invention patents and damage awards tend to be somewhat lower, but these differences are not orders of magnitude, and they are partly offset by the ease and speed with which utility models can be obtained.

Chinese invention patent applications have also seen a rapid increase in recent years, placing an internationally unprecedented workload upon the shoulders of Chinese patent examiners. China's patent examiners, on average, have far more patents to process than their American and European counterparts, and this gap is increasing over time. By 2016, Chinese examiners were evaluating 66% more patents per examiner than their American counterparts and 115% more patents per examiner than their European counterparts. It is difficult to imagine high-quality evaluation of patent applications given these burdens and the relatively low level of monthly salary Chinese patent examiners receive (anecdotal evidence suggests a salary of around 10,000 RMB, which is USD 1,600). Moreover, the recent policy reforms initiated by the central government to reduce patent filing backlogs by reducing examination time could potentially aggravate the burdens of Chinese patent examiners and thus further degrade the quality of examination.

If patents are poorly evaluated and poorly enforced, why have so many Chinese firms taken out so many patents? To answer that, we can appeal to a growing literature that documents the powerful financial incentives the Chinese government has implemented to encourage patent applications. Details can be found in our Online Appendix.

3 Model

3.1 Introduction to the Model

Having laid out some of the critical shortcomings of China's patent system, we now introduce a growth-theoretic model that illustrates their impact on Chinese growth, innovation, and welfare. We use a variant of the familiar quality ladder model (Aghion and Howitt, 1992), in which there is a continuum of differentiated intermediate goods. In each goods market at every point in time there is one firm which possesses a technological advantage over all others. In equilibrium, this market leader supplies all demand and practices limit pricing to keep the second best firm from entering the market at a lower price. However, by innovating, another potential entrant can seize technological leadership from the previous market leader and displace it in the market.

In the model, incremental and radical innovation are distinguished by their inventive step size distributions, which represent proportional improvements in productivity.[3] Once a new innovation is introduced, the technology embodied in that product quickly diffuses to follow-on innovators, enabling them to build on even a radical innovation at low cost. (This presumes a large potential number of entrants with advanced manufacturing capabilities—which likely exists in contemporary China.) The patent system itself facilitates this knowledge transfer, by requiring inventors to disclose the workings of their technologies. However, disclosure may not be balanced by the appropriate breadth of protection in a weak patent system. As soon as a follow-on innovator creates even a very minor, incremental improvement to the work of the radical innovator, this follow-on innovator can introduce a product that is slightly better, and can therefore displace the radical innovator from the market, perhaps long before the profit stream earned by the innovator has compensated her for the higher cost of radical innovation. We will show that this can lead to equilibria where firms underinvest in radical innovation.

The introduction of broad, effectively enforced patents restores the incentives for radical innovation. A market leader can still be displaced, but only by an innovator that offers a substantive improvement. The higher the bar required to constitute a patentable innovation (the breadth), the more profitable radical innovation looks, as it generates more weight in the tails of the distribution. Intuitively, technological progress becomes characterized by fewer steps, but longer strides along the quality ladder. This shift leads to an acceleration of innovation, a higher rate of economic growth, and a higher level of national welfare.

In a quality ladder model, our theoretical representation of patent breadth or scope is "vertical" in the sense that innovations lie on a line. A level of patent breadth sufficient to induce more radical or substantive innovation will feature both adequate "leading breadth" (reserving a sufficient space on the quality line ahead of the focal idea to protect it from minor improvements) and adequate "lagging breadth" (reserving a sufficient space on the quality line behind the focal idea to protect it from minor "downgrades" that limit

[3]The use of the term "radical innovation" may suggest to some readers truly revolutionary advances like Watt's steam engine or, at least, the initial introduction of the iPhone. We use the term loosely. A small number of the innovations we label "radical" will be truly revolutionary, in this sense. The majority, however, will fall short of "revolutionary," but will represent a greater inventive step than the incremental innovations with which we are comparing them.

performance but lower the price). In this, we are inspired by the taxonomy of O'Donoghue et al. (1998), and our notion of optimal patent breadth is similar in some ways to theirs.[4] Similarly to our paper, Chen et al. (2018) focus on the distinction between "safe" and "risky" innovation, and allow firms to choose among these two modes of research. This dichotomy may seem very similar to that of "incremental" and "radical", but ends up producing a somewhat different effect of patent breadth on risky/radical innovation, specifically an inverted-U shape, rather than an increasing relationship as we see in our paper. This difference arises from the fact that in our model, both types of innovation are in fact risky, radical innovation is simply more risky than incremental.

The next several pages present our formal derivation of the model. First we will solve for the static product market outcome of the model, after which we will analyze the dynamic incentives for innovation and the resulting equilibrium. Finally, we will solve the social planner's problem for this same model, characterize the nature of the inefficiency present, and consider potential policy remedies.

3.2 Product Market

There is a composite final good that is produced using a unit continuum of intermediate goods $j \in [0, 1]$. There is also a continuum of firms, indexed by f, each with a particular productivity in a given product line q_{jf}. Though it is not assumed, the equilibrium outcome will be that each intermediate good is produced by the sole producer having the highest productivity for the production of the particular line. This "productivity" can be viewed as either a marginal cost of production or a vertical measure of product quality—a "hedonic q". Given the context of our model, the latter interpretation may be more useful. We will often refer to the productivity of the state-of-the-art producer as q_j.

Each intermediate is produced linearly according to $y_j = q_j \ell_j$, where ℓ_j denotes the amount of production labor used in producing line j. They are then aggregated into a final product

[4] As we think about how to map the representation of innovation in our model to the necessarily more complicated "real world," we implicitly posit a level of patent breadth that protects a focal innovation from horizontal competition with close substitutes of equivalent quality. Our notion of adequate patent breadth therefore has an implicit horizontal dimension, as well as the explicit vertical dimension formally incorporated into our model.

according to the Cobb-Douglas style production function

$$\log(y) = \int_0^1 \log(y_j) dj$$

We assume that the intermediate goods are aggregated competitively, that is, there are many firms that can operate this technology, and competition among them will drive profits of final good aggregators to zero. Given this arrangement, intermediate firms post a price p_j and the final good aggregators decide how much to purchase conditional on that. This implies an inverse demand function for a given intermediate good j of the form

$$p_j = \frac{\partial y}{\partial y_j} = \frac{y}{y_j}$$

Of particular interest here is that the implied revenue level of an intermediate producer will be y, regardless of the price posted. Thus in the absence of competition, they would optimally like to choose an arbitrarily high price (and low quantity) in order to yield a maximal profit of $\pi_j = y$. However, they do face competition in the form of the firm with the next best productivity in producing intermediate good j, whose productivity we will refer to by q_{-j}. The "hedonic" approach suggested above reframes the productivity level of the next best producer as a quality level.

We will be particularly interested in the ratio of these productivities (or quality levels) $\lambda_j \equiv q_j/q_{-j} \geq 1$. In this case, the optimal choice of the leading producer will be to charge a price equal to the marginal cost (or quality level) of the next best producer (this is often called limit pricing). Given a wage for production workers of w, this implies

$$p_j = MC_{-j} = \frac{w}{q_{-j}} = \frac{\lambda_j w}{q_j} \quad \Rightarrow \quad y_j = \frac{q_j y}{\lambda_j w}$$

With a bit of algebra, we can derive the following outcome in terms of profits and labor costs, which also functions as a sort of income decomposition between profits and labor at the product line level

$$\tilde{\pi}_j = 1 - \lambda_j^{-1} \quad \text{and} \quad \tilde{w}\ell_j = \lambda_j^{-1}$$

where $\sim$ denotes a variable that is normalized by total output y, so for instance $\tilde{w} \equiv w/y$. Aggregating this to the economy wide level by integrating over j, we find the relations

$$\tilde{\Pi} = 1 - \Lambda^{-1} \quad \text{and} \quad \tilde{w}P = \Lambda^{-1}$$

where $\Lambda^{-1} \equiv \int_0^1 \lambda_j^{-1} dj$ and $P \in [0,1]$ is the share of labor devoted to production. The above relation between the normalized wage $\tilde{w}$ and production labor P will prove useful in solving for the model equilibrium. The final piece of the puzzle can be obtained by plugging the expression for y_j above back into the final good production function. In conjunction with the wage equation, this yields

$$y = PQ \cdot \frac{\Lambda}{\Omega}$$

where we must additionally define the following aggregates $\log(Q) \equiv \int_0^1 \log(q_j)dj$ and $\log(\Omega) \equiv \int_0^1 \log(\lambda_j)dj$. Notice that in the case where there is no heterogeneity in terms of λ_j, this reduces to $y = PQ$. In any case, we must have $y \leq PQ$ since $\Lambda \leq \Omega$ as show below using Jensen's inequality

$$\log(\Omega^{-1}) = \int_0^1 \log(\lambda_j^{-1})dj \leq \log\left(\int_0^1 \lambda_j^{-1} dj\right) = \log(\Lambda^{-1})$$

In this way, the ratio Λ/Ω can be seen as a measure of labor misallocation across product lines induced by markup heterogeneity.

3.3 Innovation

Innovation is the process by which productivity (or quality) gains are realized. There is a large pool of potential entrants that employ researchers in hopes of innovating. Upon successfully innovating, they increment the state-of-the-art productivity (or quality) for a randomly chosen product line j by some random factor.

There are two types of innovation: incremental (i) and radical (r), where will use $k \in \{i, r\}$ to denote a generic type. Each type of innovation has a certain per-researcher-skill probability of success η_k and a distribution of realized innovative step sizes $\gamma > 1$ given by a Pareto distribution based at 1, which has distribution functions

$$F_k(\gamma) = 1 - \gamma^{-\alpha_k} \quad \Rightarrow \quad f_k(\gamma) = \alpha_k \gamma^{-\alpha_k - 1}$$

It is natural and without loss of generality to assume that $\alpha_i > \alpha_r$, meaning the distribution of innovative step sizes associated with radical innovation has a higher mean and thicker tails. This situation is depicted visually in Figure 2, where we plot the respective density functions and the ratio of survival functions (the probability of exceeding a certain level) between radical and incremental innovations.[5]

The notion of a patent here affords a firm protection for their technology, as well as protection against technologies that are sufficiently similar or insufficiently novel. We operationalize this formally by saying that the patent extends to technologies with productivity Bq_j or less, where we refer to $B > 1$ as the patent breadth. Given that new innovations build upon the current state-of-the-art q_j, this means innovators must achieve a step size of at least B to be viable given the incumbent's patent protection. In order for patents to protect a space of breadth B, it must be the case that patent infringement damage awards are large enough to deter entry by a potential innovator. In Appendix B.1 we show that the expected damage award has to exceed the expected profit that would accrue to any minor innovator whose efforts would land within B units of the focal invention, and that this damage award is increasing in B.

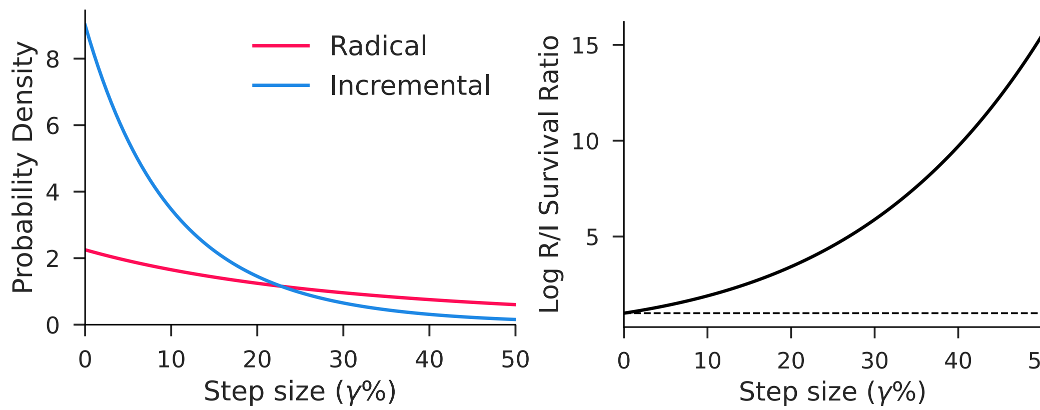
On the labor market side, there is a common pool of workers of mass 1 that can either engage in production or research. However, though each worker has common aptitude in production, they are heterogeneous in their aptitude for research. In particular, worker $i \in [0, 1]$ has research aptitude of $(1-\beta)i^{-\beta}$. Thus, if the total research share is R, assuming

[5]To analyze the various expectations and probabilities in the derivations that follow, it will be useful to know the following facts about Pareto distributions

$$\mathbb{E}_\alpha\left[\frac{B}{\gamma}\Big|\gamma > B\right] = \frac{\alpha}{1+\alpha} \qquad \mathbb{E}_\alpha\left[\log\left(\frac{\gamma}{B}\right)\Big|\gamma > B\right] = \frac{1}{\alpha} \qquad \mathbb{P}_\alpha[\gamma > B] = B^{-\alpha}$$

Note that the former two equations are essentially statements that the Pareto distribution is memoryless in its conditional expectation.

Figure 2: Incremental and Radical Innovation Step Size Distributions



this constitutes the most productive researchers, then the effective amount of research undertaken is

$$A \equiv \int_0^R (1-\beta) i^{-\beta} di = R^{1-\beta}$$

We assume that firms cannot observe researcher type, so they expect to get the average research aptitude $a(R) \equiv A/R = R^{-\beta}$. The net effect of this arrangement, compared to one with homogeneous workers, is that there are decreasing returns to research in the aggregate, as higher research intensity leads firms to employ less and less skilled researchers. Note that we still assume that the current stock of knowledge enters linearly into the production of new ideas, meaning our model would feature scale effects in the overall population size (which we assume to be fixed at 1) as demonstrated in Jones (1995).

Finally, the government can also subsidize innovation costs. We denote this subsidy rate by $s \in [0, 1]$. The costs of this subsidy program is funded by a lump-sum tax on consumers that balances the government's budget at each moment in time.

3.4 Equilibrium

Now we can derive the value of obtaining a new product line for both incremental and radical innovation. In general, the value of a product line j satisfies

$$r v_j - \dot{v}_j = \pi_j - \tau v_j$$

where τ is the combined rate of creative destruction from both incremental and radical innovation, which we take as given for now. Letting $g_v \equiv \dot{v}_j / v_j$, this leads to the equation

$$v_j = \frac{\pi_j}{r - g_v + \tau}$$

We will restrict attention to steady state outcomes. In this case, note that we will have $g_v = g_\pi = g_y = g_Q \equiv g$. Under the assumption of log utility on the part of consumers, the Euler equation is $r = \rho + g$. Since the normalized value $\tilde{v}_j$ depends only on λ_j via $\tilde{\pi}_j$, we can write it as a generic function of λ rather than a specific j

$$\tilde{v}(\lambda) = \frac{\tilde{\pi}(\lambda)}{\rho + \tau} = \frac{1 - \lambda^{-1}}{\rho + \tau}$$

To determine the equilibrium level of innovation, we must calculate the expected present value of successful innovation. This will result from the combination of the probability of an innovation large enough to exceed B and the expected profits conditional on that happening. Thus the value of innovation of type k being given by

$$\tilde{z}_k = \int_B^\infty \tilde{v}\left(\frac{\gamma}{B}\right) dF_k(\gamma) = \frac{1}{1 + \alpha_k} \frac{P_k}{\rho + \tau}$$

where $P_k = \mathbb{P}_k[\gamma > B] = B^{-\alpha_k}$ is the probability of an innovation of a given type exceeding the patent novelty threshold B. Given an overall research share of R, the innovation rate of research type k will be $\tau_k = R^{-\beta} \eta_k R_k$. This then leads to the aggregate innovation rate $\tau = \tau_i P_i + \tau_r P_r$. In the case where there are positive amounts of type-k research in equilibrium, we will have the following free entry condition

$$R^{-\beta} \eta_k \tilde{z}_k = (1 - s)\tilde{w} \tag{1}$$

Since firms can choose costlessly between incremental and radical innovation, we will in general only see one type of innovation in equilibrium, and this will depend on whichever type maximizes the quantity $\eta_k \tilde{z}_k$. We now arrive at an equation characterizing the equilibrium value of research labor R_k for the case when type k innovation is predominant

$$\frac{1}{1+\alpha_k}\frac{\eta_k R_k^{-\beta} P_k}{\rho+\eta_k R_k^{1-\beta} P_k} = \frac{(1-s)\Lambda^{-1}}{1-R_k} \tag{2}$$

Since creative destruction at the product line level occurs at rate τ, regardless of product characteristics, the distribution over λ_j will be identical to that of the incoming distribution F_k, and so we will have $\Lambda^{-1} = \frac{\alpha_k}{1+\alpha_k}$. Thus we arrive at the simplified equation

$$\frac{\eta_k R_k^{-\beta} B^{-\alpha_k}}{\rho+\eta_k R_k^{1-\beta} B^{-\alpha_k}} = \frac{(1-s)\alpha_k}{1-R_k}$$

It is clear due to monotonicity and Inada-like conditions that this equation has a unique and interior solution in R_k. Furthermore, the equilibrium value of R_k will be decreasing in B. The growth rate resulting from this research share, accounting for researcher aptitude and patent policy B will be

$$g_k = \left[\frac{1}{\alpha_k} + \log(B)\right] \cdot \eta_k B^{-\alpha_k} \cdot R_k^{1-\beta} \tag{3}$$

This specifies the equilibrium allocation of labor towards research conditional on innovation of type k prevailing. Exactly which type of innovation prevails in equilibrium will depend on the relative values of $\eta_k \tilde{z}_k$, which reduces to the comparison

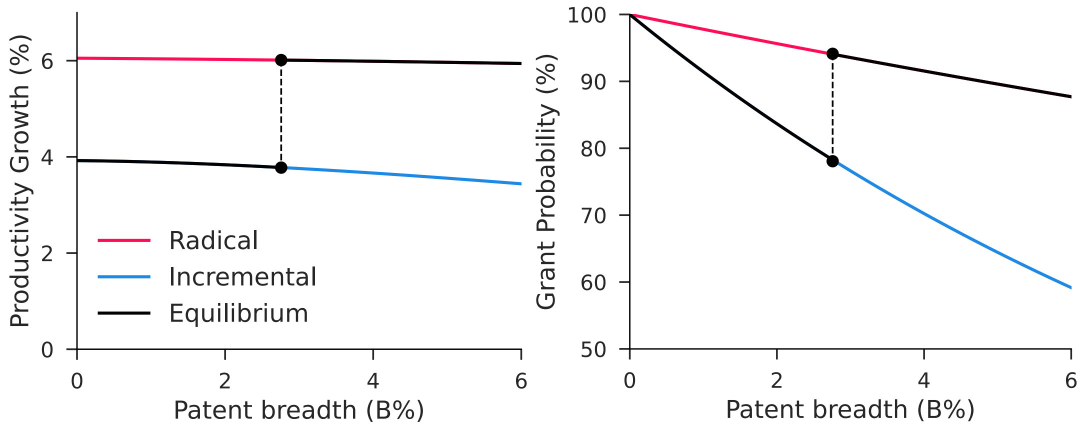
$$\frac{\eta_i B^{-\alpha_i}}{1+\alpha_i} \quad \lesseqgtr \quad \frac{\eta_r B^{-\alpha_r}}{1+\alpha_r}$$

3.5 Effect of Policy

Because $\alpha_i > \alpha_r$, an increase in the patent breadth B will have a greater effect on the incentives for incremental innovation, as its thinner tails make it less likely to exceed a given threshold. Generally, we will also assume that in the absence of any appreciable patent breadth ($B = 1$), incremental innovation prevails, meaning

$$\frac{\eta_i}{1+\alpha_i} > \frac{\eta_r}{1+\alpha_r}$$

Figure 3: Output Growth and Grant Probability as a Function of Patent Breadth



Note: *Lines denoted "Radical" and "Incremental" are conditional rates, while the line denoted "Equilibrium" is that which prevails in equilibrium.*

In this case, there is a specific B^* above which radical innovation becomes dominant, and this is given by

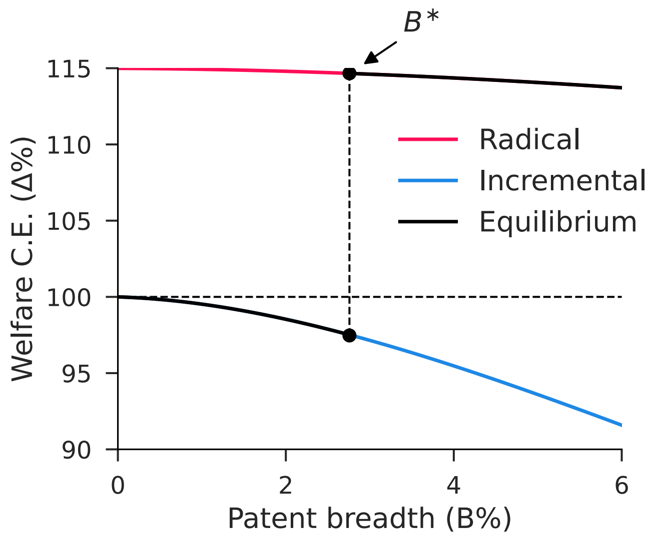
$$B^* = \left[\frac{\eta_i}{\eta_r}\frac{1+\alpha_r}{1+\alpha_i}\right]^{\frac{1}{\alpha_i-\alpha_r}} > 1$$

However, it is important to note that, conditional on type k, increasing B depresses both the amount of labor allocated to innovation and the rate of successful innovation, as some (small) innovations are effectively thrown out. Thus a fully optimal policy will have to manage both the margin between radical and incremental innovation and the overall innovation allocation.

It is useful at this point to visualize the effect of patent breadth on particular outcomes such as the growth rate and overall welfare. For the following figures we use $\alpha_i = 9.01$, $\alpha_r = 2.25$, $\eta_i = 0.80$, and $\eta_r = 0.22$. As can be seen in Figure 3, this ensures that with no patent breadth ($B = 1$) incremental innovation prevails, and that past a certain threshold (B^*) there is a discrete switch to radical innovation. This transition to radical innovation yields large gains in terms of output growth, though it comes as the cost of throwing away some innovations that fall below B, as can be seen by the fact that the type-conditional growth values are decreasing in B.

To understand the net effect of these changes, in Figure 4 we plot the consumption equiva-

Figure 4: Total Welfare as a Function of Patent Breadth



Note: *Values denote consumption equivalents of welfare (W_1) relative to equilibrium welfare when $B = 1$ (W_0), calculated using CE $= 100 \times \exp(W_1 - W_0)$.*

lent welfare levels corresponding to the various outcomes as a function of patent breadth B. In addition to changes in aggregate growth rates, this accounts for changes in overall output arising from changes in production labor levels and the misallocation of input factors due to monopoly distortions.

3.6 Social Planner

Now let's solve the social planner's problem. Given our distributional assumptions, the growth rate of the aggregate productivity index Q (and hence steady state output Y) can be expressed as

$$g = \tau_i \mathbb{E}_i[\log(\gamma)] + \tau_r \mathbb{E}_r[\log(\gamma)] = \frac{\tau_i}{\alpha_i} + \frac{\tau_r}{\alpha_r}$$

At this point, we can construct the Hamiltonian corresponding to the social planner's optimization problem. This amounts to choosing R_i and R_r, which determines the overall growth rate g and the level production via $P = 1 - R_i - R_r$. Note that this already builds in an efficient allocation of production labor wherein an equal quantity of labor $\ell_j = P$ is devoted to each product line, and hence $Y = QP$. The Hamiltonian is then

$$H = \log(Q) + \log(1 - R_i - R_r) + \mu Q(R_i + R_r)^{-\beta} \left(\frac{\eta_i R_i}{\alpha_i} + \frac{\eta_r R_r}{\alpha_r} \right)$$

Given the frictionless choice between incremental and radical innovation, we would once again expect a so-called "bang-bang" result, whereby only one form of innovation arises at the optimum and the transition between the two occurs abruptly at some cutoff value. And so the determination of which form of innovation is actually dominant at the optimum will depend upon the relative values below

$$\frac{\eta_i}{\alpha_i} \quad \lesseqgtr \quad \frac{\eta_r}{\alpha_r}$$

In this case, the first order condition corresponding to the non-dominant form of innovation will be strictly negative, while that of the dominant form will hold with equality. Conditional on some dominant form of innovation k, we then arrive at optimality conditions

$$\begin{aligned} 0 = H_{R_k} &= -\frac{1}{P} + \mu Q(1 - \beta) R_k^{-\beta} \frac{\eta_k}{\alpha_k} \\ \rho\mu - \dot{\mu} = H_Q &= \frac{1}{Q} + \mu g \end{aligned}$$

Finally, we can manipulate these equations so as to eliminate the costate variable μ. In steady state, regardless of the dominant form of innovation, we should have

$$\frac{\eta_k}{\rho} \frac{R_k^{-\beta}}{1 - \beta} = \frac{\alpha_k}{1 - R_k}$$

and the resulting growth rate is $g_k = \frac{\eta_k}{\alpha_k} R_k^{1-\beta}$.

3.7 Optimal Policy

Given knowledge of the true model parameters, one can induce radical innovation with minimal loss of ideas by setting exactly $B = B^*$. Thus one approach to optimal policy is to set $B = B^*$ when it is optimal to induce radical and $B = 1$ otherwise. Whether inducing radical is optimal will depend on whether the gains from switching to the higher step size distribution are outweighed by the additional research costs and the need to "throw away" innovations with step size falling below B. This is an interesting objective, but it should be noted that although it will ensure the "right" type of innovation occurs, it doesn't necessarily get the overall split between production and innovation right.

That said, one could couple this with an overall subsidy on research to try and achieve something closer to a full optimum. It is straightforward to show that in the presence of a uniform (across radical and incremental) research subsidy, the equilibrium research allocation is a monotone increasing function of s ranging from zero to one, hence there is a unique $\hat{s}$ that yields the socially optimal level of research, which will satisfy $R_k^*(\hat{s}, B^*) = \hat{R}_k$. Additionally, since the research subsidy is uniform, it doesn't further distort the margin between incremental and radical innovation, so employing the B^* policy still ensures that the optimal type of innovation prevails.

It may be useful to reflect on how the research subsidy modeled here is reflected in real world policies. Here we think of the subsidy as investment in education and basic R&D, both of which enable (some) workers to become researchers. For a country at its level of income per capita, China has arguably done an impressive job of investing in the human capital of its urban workforce and the establishment of scientific research capabilities in its leading research universities. Rozelle and Hell (2020) point to much lower levels of investment in the education and health of the rural population as a cause of concern, but note that increasing these investments to close the education/health gap with urban populations lies well within the fiscal capacity of the Chinese state. If that goal were met, China might be close to the optimal level of "research subsidy" featured in our model.

The only remaining source of inefficiency is the labor misallocation induced by monopoly distortions. Neutralizing these would require a product line specific labor subsidy that depends on that line's realized step size λ_j. Supposing this policy is employed, one can then use the same B^* and a modified $\hat{s}$ to implement the first best optimum.

4 Patent Fee Based Policy

Now consider an alternative policy in which there is no explicit novelty threshold B, but firms must pay a fee d in terms of final good to acquire a patent. Since granting is not stochastic conditional on innovation size, this can be seen as either an application fee or a granting fee. As the expected present value of an innovation is increasing in innovation size (γ), a fixed fee will induce a cutoff above which firms will choose to file a patent. We will also call this cutoff B, as it will operate similarly, but we will construct a mapping between d and B.

Because total output is growing continually, we must index d to output to achieve stationarity. Thus we will define $d = \tilde{d}Y$ and think of the indexed $\tilde{d}$ as the primary variable chosen by the policymaker. Now it is straightforward to see that a firm will only file for a patent if $\tilde{v}(\gamma) \geq \tilde{d}$, and so the cutoff value B will satisfy

$$\frac{1 - B^{-1}}{\rho + \tau} = \tilde{v}(B) = \tilde{d}$$

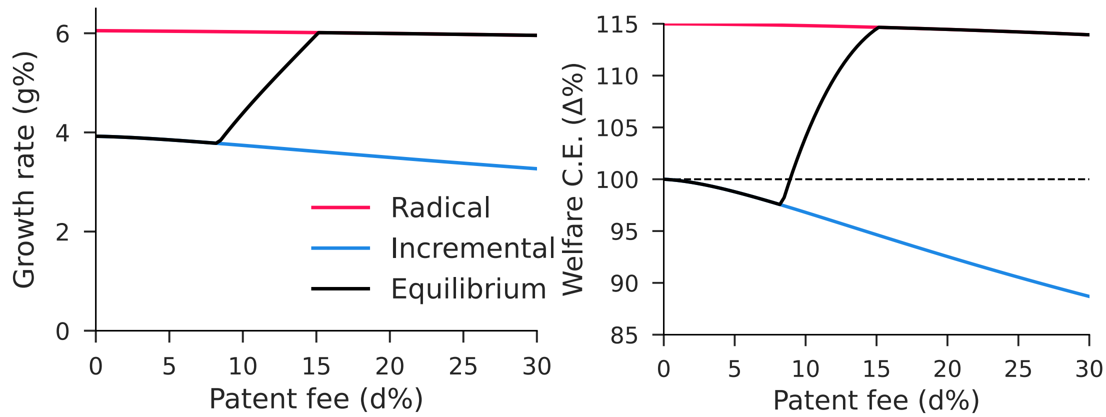
If $\tilde{d}$ too large, it will preclude any patenting. Thus we will assume it is small enough for a finite B and verify this condition ex post. In that case, we find that B is implicitly (as τ depends on B itself) given by

$$B = \frac{1}{1 - \tilde{d}(\rho + \tau)}$$

The full derivation can be found in Appendix B.2, but the main takeaway is that, contrary to the baseline case, this type of policy induces a smooth transition between incremental and radical innovation regimes. For sufficiently large or small $\tilde{d}$, either incremental or radical innovation will prevail. In this case, we can show that the equation characterizing the equilibrium is identical to the B policy case. However, there will be a range of $\tilde{d}$ values for which the implied B will be constant at B^* and firms will collectively undertake a mixture of incremental and radical innovation. This arises because when firms switch between incremental and radical innovation, the change in τ changes the linkage between B and $\tilde{d}$.

Letting τ_i^* and τ_r^* be the respective innovation rates for the case where $B = B^*$, these pure

Figure 5: Growth and Welfare as a Function of Patent Fee



regimes will arise when

$$\tilde{d} \leq \tilde{d}_0 \equiv \frac{1 - 1/B^*}{\rho + \tau_i^*} \quad \text{or} \quad \tilde{d} \geq \tilde{d}_1 \equiv \frac{1 - 1/B^*}{\rho + \tau_r^*}$$

For $\tilde{d} \in (\tilde{d}_0, \tilde{d}_1)$, the effective B will remain constant at B^*, meaning

$$\tau = \frac{1 - 1/B^*}{\tilde{d}} - \rho$$

and there will be an interior outcome, with some firms engaging in incremental innovation and others engaging in radical innovation. The effect of patent fee policy on aggregate growth and welfare using our estimated parameters is depicted in Figure 5.

4.1 Additional Extensions

In the Online Appendix, we cover three additional extensions. The first (OA.I) allows for the possibility of imperfect adjudication of patent novelty, which reflects the fact that examiners are not simply given a numerical step size (γ) but must estimate this based on a technical description of the invention. The second (OA.II) considers the case where innovation costs (η_k) are not fixed but are instead drawn from some joint distribution. This yields outcomes that are continuous in the patent breadth, as opposed to the abrupt transition that we see in the baseline model. The third (OA.III) replaces incremental innovation with a deterministic outcome so as to isolate the risk channel yielding very similar results to the baseline model.

5 Data and Calibration

In this section, we briefly discuss the data and strategy we use to calibrate key parameters of the growth model we introduced in the previous section.

5.1 Data Sources

We compile a dataset containing all Chinese invention patents filed by domestic Chinese firms from 2001 to 2008 [6]. This time period gives us sufficient time to observe the full life cycle of a patent (e.g. citations to the patent), thus mitigating truncation problems. Restricting to this time period, which is between two major patent reforms, also keeps the regulatory environment stable across the sample. We collect the data from IncoPat, a private data provider widely used by scholars and Chinese patent examiners. Our sample contains 738,974 invention patents filed by 40,691 domestic Chinese firms.

We identify unique inventors based on the firm they work with, the distribution of technological areas they filed patents in, and their names [7]. This procedure yields 340,724 unique Chinese inventors. We measure their productivity using the number of patents they have filed and the forward citations received by these patents within a 10-year horizon (from the patent filing year to the year the patent was cited). Additionally, we count the number of years in which an inventor filed patents to control for entry into and exit from the sample.

5.2 Identification

In the baseline model, we assume that the enforced patent breadth is minimal ($B = 1$) and that parameters are such that incremental innovation prevails in equilibrium. As such, we can only identify incremental research production parameters (α_i and η_i) from the data, in addition to general parameters such as the research aptitude distribution (β). Below we provide some intuition for how we calibrate these parameters.

[6] For the forward citations in Table 6, we extend the data window to the year 2013.

[7] Chinese names usually contain two or three characters without any middle names. It is not uncommon for many Chinese people to share the same name. By controlling the firm inventors patent with and the technological areas they mainly invent in, it is easier for us to identify unique inventors via their names.

Research Aptitude Distribution. To calibrate the parameter controlling the distribution of research aptitude across the populace (β), we need to gauge the research productivity of a group of active inventors in China. Previous studies (Azoulay et al., 2010, 2019) look at the research productivity of top academic scholars, commonly measured by the number of citations to their papers. It is typically a complicated process for academic knowledge to be transferred to specific applications and utilized by firms (Jensen and Thursby, 2001; Thursby and Thursby, 2002). We instead look at the productivity of active inventors employed by domestic Chinese firms.

Given data on the productivity of inventors in terms of number of patents per year, we can use these to construct an empirical cumulative distribution function (CDF), which we can call G. Since worker i has research aptitude $a(i) = (1-\beta)i^{-\beta}$, this resuls in CDF of

$$G(a) = 1 - a^{-1}(a) = 1 - \left(\frac{a}{1-\beta}\right)^{-\frac{1}{\beta}}$$

By looking at the log of the survival function (one minus the CDF), we arrive at an equation amenable to linear regression

$$\log\left(1 - G(a)\right) = \frac{1}{\beta}\log(1-\beta) - \frac{1}{\beta}\log(a)$$

The non-intercept coefficient of this regression yields an estimate of $1/\beta$, which we use directly as one of our moments and can clearly yield information on β itself.

Step Size Distributions. The parameter α_i governs the distribution of inventive step sizes for incremental innovation. Several metrics are used by the literature to measure innovation, most of which are based on citations and technological classifications of patented inventions (March, 1991; Fitzgerald et al., 2021). We use total forward citations as our main metric of the innovation step size for a given patent.

Of course, we can't simply equate citation counts with the inventive step size, as they have different scales and one is a continuous variable while the other is discrete. Thus we employ an auxiliary model to describe the distribution of citation counts (c) conditional on a particular step size (γ) for some scaling factor $\kappa > 0$, namely $\mathbb{E}[c|\gamma] = \mathbb{V}[c|\gamma] = \kappa(\gamma - 1)$, which would be consistent with a Poisson distribution. In this case, integrating out γ, which is Pareto distributed, yields unconditional expectations

$$\mathbb{E}[c] = \frac{\kappa}{\alpha - 1} \quad \text{and} \quad \mathbb{V}[c] = \frac{\kappa}{\alpha - 1} + \frac{\kappa^2 \alpha}{(\alpha - 1)^2 (\alpha - 2)}$$

The full derivation for these equations can be found in Appendix B.3. Using these, we can construct an index of dispersion that depends only on α but not κ

$$D \equiv \frac{\mathbb{V}[c] - \mathbb{E}[c]}{\mathbb{E}[c]^2} = \frac{\alpha}{\alpha - 2}$$

For a pure Poisson process, this dispersion parameter would be zero. The overdispersion caused by variation in γ causes it to be larger. Thus we can cleanly identify this parameter from citation data in the case that $\alpha > 2$.

Innovation Cost Parameters. Innovation cost in our model corresponds to a value for η_i, which is actually an innovation rate and so inverse cost. In general, innovation cost parameters influence both the growth rate and the fraction of labor devoted to research. Higher costs directly mediate the linkage between the research share and the aggregate growth rate. In addition, higher costs will influence the research share directly via the incentives for innovation.

5.3 Calibration Results

We calibrate the model by targeting a set of moments that are informative about the underlying parameters. Though we will give intuition about how and why the calibration works, the system is not diagonal and each moment can be influenced by a number of parameters. Further, because we do not have standard errors for some of the moments, we cannot produce standard errors for the estimated parameters, though we do provide certain robustness checks.

In Table 3 we list the moments used for the calibration along with their source, data values, and predicted values at optimum. The estimated parameters are listed in Table 4. Note that we are implicitly assuming that innovation is entirely incremental in the baseline, and

Table 3: Moments Used for Calibration

Moment Name	Data	Model
Aggregate researcher share (R)	0.10	0.103
Aggregate TFP growth (g)	0.04	0.039
Citation coefficient of variation ($\mathbb{C}$)	1.30	1.285
Inventor distribution tail index ($1/\beta$)	1.56	1.560

Table 4: Estimated Parameter Values

Parameter Name	Value
Discount rate — fixed (ρ)	0.050
Research aptitude distribution (β)	0.641
Incremental step size distribution (α_i)	9.009
Incremental innovation rate (η_i)	0.800

so we only have estimates for general parameters and incremental parameters (α_i and η_i). We have a separate calibration for the radical parameters and also entertain a wide variety of parameter values when considering radical innovation.

5.4 Radical Innovation

Step Size Distribution. The scenario we consider is special in the sense that we need to evaluate radical innovations in an economy that is still transitioning to the technological frontier (China), as opposed to one the is operating at the frontier and extending it (e.g. the US). That said, we need a group of inventions whose inventive steps could be measured in both systems. To this end, family patents, a group of patents protecting the same invention filed in a variety of jurisdictions, might be a useful phenomenon here.

Specifically, we use the aforementioned data set on Chinese patents. We identify the group of Chinese patents made by domestic Chinese firms that were also registered as patents in the U.S. (i.e. family patents). We use these Chinese patents with U.S. counterparts as a proxy for radical innovations and label all the Chinese patents without any foreign family patents as incremental innovations. While this metric for incremental versus radical innovations might not be suitable for patent systems with relatively high volumes of family patents such as US or the EU, it may be suitable in the context of China where very few patents attain such a status (see Section 2).

We use the number of forward citations received by the focal patent within 10 years of patent filing as the measure for inventive step size. Table 5 lists the distribution of 10-year citations for several groups of patents we use in our calibration [8]. Row 1 and 4 are the same group of patents, Chinese patents made by domestic Chinese firms that are also filed with foreign patent offices, measured by forward citations received in the US and forward citations received within China, respectively. The US patent data is collected from PatentsView. Thus, rows 1 and 4 serve as a "bridge" connecting our measure for the inventive step size of Chinese patents with the inventive step sizes of patents registered in the US. The distribution of forward citations to the population of the US patents is listed in row 2. Row 3 reports the forward citations to US patents with self-citations excluded[9]. Row 5 reports the distribution of forward citations to Chinese patents made by domestic Chinese firms that have not been registered as patents in the US, the proxy for incremental innovations in China.

Innovation Cost Parameters Calibrating the cost by innovation type, however, is complicated due to the difficulty of clearly distinguishing different types of innovation. For instance, firms might invest a lump-sum fixed cost before any meaningful R&D activities begin, such as the cost of hiring researchers and acquiring equipment. These costs, in theory, need to be spread across all innovation outputs the firm produces later, regardless of the type. With these complexities on mind, we implement two approaches and exploit another two data sets to calibrate the innovation cost, attempting to cross-validate the calibration for this parameter.

To start with, we compile a database comprising Chinese manufacturing companies, which tend to have fewer diversified product lines and smaller sizes. We begin with the Chinese Annual Survey of Manufacturing (ASM), a longitude survey conducted by the Bureau of Statistics which is widely used by scholars. After matching these firms with their

[8] Only patents with positive citations are included.

[9] We do not report the self-citations-excluded measures for Chinese patents because, contrary to the US, Chinese patents have very limited number of citations coming from the same patentee.

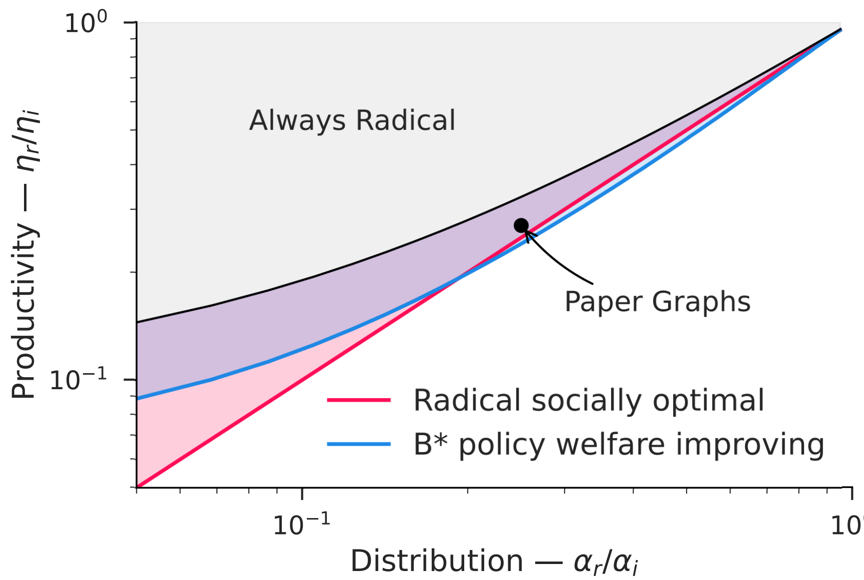
patent records, we regress the proxies for radical and incremental innovations on firms' R&D spending. We relegate the details to Appendix E. Table 7 reports the results. As a robustness check, we investigate different transformations of these variables: standardized variables in columns 1 and 2; logarithm transformed variables in columns 3 and 4; raw numbers in columns 5 and 6. These results indicate that the ratio of family patents—a proxy for radical innovations—to non-family patents—a proxy for incremental innovations—is around 2.12.

One might be concerned that the data we use for this approach might only reflect the situation 10 years ago or more. Therefore, we try an alternative approach on a recent sample of all publicly traded companies listed in mainland China which have records of R&D activities for at least two consecutive years. In addition to regressing patent output on R&D spending, as we have done in Table 7, we also regress R&D spending on the proxies for radical innovations and the number of incremental innovations. We relegate details on specifications and results to Appendix E. The estimates from this approach indicate that the ratio of the average cost of one radical innovation to that of an incremental innovation ranges from 15.8 to 21.3. While these ratios are higher than our previous estimates, they imply an alternative scenario in which radical innovation is much more expensive relative to incremental innovation. The takeaway here is that we do not stick to any specific parameter calibration for the radical innovation parameters. We simply explore possible scenarios using different sets of parameters. Indeed, it is unlikely that a fast-changing innovation system like the one in China "sticks to" a particular set of parameters over time.

5.5 Sensitivity Analysis

Due to the substantial uncertainty surrounding the exact values of the radical innovation parameters α_r and η_r, we attempt to provide a comprehensive picture of the efficiency properties and policy implications for a wide range of values. To this end, we construct a phase diagram delineating the various segments of this space. In Figure 6, we plot the relative values of α_r and η_i in logarithmic space. Shaded in gray, there is a large region of η_r values for which radical innovation is always the equilibrium outcome, even when

Figure 6: Sensitivity of Breadth Policy Outcome to Radical Parameters



$B = 1$, which in essence means $B^* = 1$. This is relatively uninteresting and we assume this is not the case throughout. More interesting are the areas shaded in color. The red area corresponds to values for which a social planner would choose radical innovation, while the blue area corresponds to values for which using the implies B^* policy would lead to welfare improvements.

Interestingly, while these two areas unsurprisingly share a large amount of overlap, there are exceptions to this in either direction. There are regions where radical innovation is in principle socially optimal, but actually achieving it with policy would involve setting a very large B^* and ultimately throwing away too many good innovations. Conversely, there are regions where radical innovation is not socially optimal but implementing B^* nonetheless leads to a welfare improvement. This arises because switching to radical innovation leads to a substantial reduction in the rate of creative destruction (though with larger step sizes), thus alleviating a business stealing inefficiency present in equilibrium. This could be partially corrected if one were to also implement an optimal subsidy in addition to the optimal breadth policy.

Finally, one might be interested in the nature of the jointly optimal patent breadth (B) and subsidy (s) policy. This becomes harder to visualize in two dimensions for a range of parameter values. In general, we find that there are parameterizations where one may wish to use only a subsidy and not induce radical innovation with a patent breadth $B > 1$. However, this subsidy is usually quite large, in the range of 60%, while the optimal subsidy once the switch to radical innovation is made falls to roughly 25%. This arises because of the aforementioned business stealing effect is much weaker with radical innovation, which has lower rates of creative destruction. Thus if one is constrained to set relatively low subsidy values, radical innovation is almost always optimal to induce.

6 Policy Implications and Conclusions

Achieving optimal patent breadth. Incentives for substantive innovation are likely to remain weak while infringement damages are low. A substantial increase in damage infringement fees for invention patents would strengthen infringement deterrence, and, therefore, the incentives for more substantive innovation. Maintaining current low levels of damage awards for utility models would tilt incentives for patent applicants against utility models and in favor of invention patents, whose quality evaluation should be substantially strengthened in the manner described below. Much higher infringement damage awards will make it more important for courts to correctly adjudicate complicated patent cases. Changes in legal regulations that encourage courts to employ expert witnesses and that permit foreign experts, foreign law firms, and evidence from related patent litigation in foreign courts to be used in reaching appropriate conclusions are likely to accelerate the pace at which Chinese courts can acquire the ability to make these difficult judgments accurately and fairly.

Promoting more rigorous patent examination and increasing patent fees. The ability of the Chinese patent office to evaluate patent quality will be increased if the number of low-quality applications is reduced and the resources devoted to patent evaluation are increased. Recent government restrictions on direct patent subsidies represent a useful step in this direction (see the online Appendix for details on China's patent subsidies). The flow of low-quality patent applications will be further reduced if the incentives to patent embodied in China's current tax code are recalibrated to reward quality rather than quantity.

The ending of government directives that set formal or informal patent application and grant targets for regions, industries, or firms could further weaken incentives to file low-quality patent applications. A significant increase in invention patent application and renewal fees would have a similar effect; the revenues generated by these (much) higher fees could increase the resources within the patent office to evaluate patent applications, raising the quality of patent examinations. Similarly, a significant increase in utility model application and renewal fees might further limit future growth in utility model applications. Requiring Chinese patent applicants to choose at the time of initial application between a utility model or an invention patent will eliminate current incentives for firms to take out multiple utility models for every invention patent.

A well-resourced post-grant review system could play an important role in the winnowing of China's existing "thicket" of low-quality patents and in the reinforcement of higher standards for inventive novelty throughout the patent system. This creates an argument for the direction of some portion of the additional revenues generated by substantially higher application and renewal fees were directed to support the expansion and acceleration of China's post-grant patent review system. New legislation that imposed a relatively high burden of proof on the parties in this post-grant review procedure seeking to defend the validity of their patents could further shift the costs generated by this system onto the patent-holder defendants and limit their incentives to defend their weakest patents.

6.1 Concluding Remarks

China has grown (very) rapidly for four decades, despite a weak patent system. This is not a contradiction of the model presented in this paper, merely a repeat of longstanding East Asian development trends. Like its East Asian predecessors on the pathway to rapid industrialization, Japan, Taiwan, and South Korea, mainland China has been able to grow rapidly through the absorption, application, and modification of foreign ideas. Like China, these areas grew fastest when their patent systems were relatively weak. This was not because weak patents are optimal in the long run, but because they were a limited impediment to growth when growth was driven primarily by imitation/absorption of foreign technology rather than indigenous innovation. But like those onetime Asian tigers, China will approach the exhaustion of opportunities for growth by catch-up, and will need to increasingly generate more substantive innovations of its own.

The model presented in the current paper has clear policy implications, providing important theoretical and empirical guidance to a longstanding debate over the shortcomings of Chinese IP policy that has often been long on vitriol but short on specifics. The model suggests that China will most effectively reach its innovative potential with fewer but better patents, (much) higher damage awards, higher patent fees, and greater discrimination between invention patents and utility models. Will these reforms be adopted? Only time will tell.

References

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr (2018): "Innovation, reallocation, and growth," *American Economic Review*, 108, 3450–91. 5

Acemoglu, Daron, Ufuk Akcigit, and Murat Alp Celik (2020): "Radical and Incremental Innovation: The Roles of Firms, Managers and Innovators," *AEJ Macroecon (forthcoming)*. 5

Agarwal, Rajshree, Martin Ganco, and Rosemarie H Ziedonis (2009): "Reputations for toughness in patent enforcement: Implications for knowledge spillovers via inventor mobility," *Strategic Management Journal*, 30, 1349–1374. 4

Aghion, Philippe and Peter Howitt (1992): "A model of growth through creative destruction," *Econometrica*, 60, 323–351. 4, 10

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde (2021): "Back to basics: Basic research spillovers, innovation policy, and growth," *The Review of Economic Studies*, 88, 1–43. 5

Akcigit, Ufuk and William R Kerr (2018): "Growth through heterogeneous innovations," *Journal of Political Economy*, 126, 1374–1443. 5

Arora, Ashish, Andrea Fosfuri, and Alfonso Gambardella (2004): *Markets for technology: The economics of innovation and corporate strategy*, MIT press. 4

Arora, Ashish and Alfonso Gambardella (1994): "The changing technology of technological change: general and abstract knowledge and the division of innovative labour," *Research policy*, 23, 523–532. 4

Azoulay, Pierre, Christian Fons-Rosen, and Joshua S Graff Zivin (2019): "Does science advance one funeral at a time?" *American Economic Review*, 109, 2889–2920. 26

Azoulay, Pierre, Joshua S Graff Zivin, and Jialan Wang (2010): "Superstar extinction," *The Quarterly Journal of Economics*, 125, 549–589. 26

Barry, Chris, Ronen Arad, Landan Ansell, Meredith Cartier, and HyeYun Lee (2017): "2017 Patent Litigation Study: change on the horizon?" . 8

Belderbos, René, Jinhyuck Park, and Martin Carree (2021): "Do R&D investments in weak IPR countries destroy market value? The role of internal linkages," *Strategic Management Journal*, 42, 1401–1431. 4

Bessen, James and Michael J Meurer (2005): "Lessons for patent policy from empirical research on patent litigation," *Lewis & Clark L. Rev.*, 9, 1. 9

Beukel, Karin and Minyuan Zhao (2018): "IP litigation is local, but those who litigate are global," *Journal of International Business Policy*, 1, 53–70. 4

Branstetter, Lee G, Raymond Fisman, and C Fritz Foley (2006): "Do stronger intellectual property rights increase international technology transfer? Empirical evidence from US firm-level panel data," *The Quarterly Journal of Economics*, 121, 321–349. 4

Branstetter, Lee G and Mariko Sakakibara (2002): "When do research consortia work well and why? Evidence from Japanese panel data," *American Economic Review*, 92, 143–159. 4

Chen, Yongmin, Shiyuan Pan, and Tianle Zhang (2018): "Patentability, R&D direction, and cumulative innovation," *International Economic Review*, 59, 1969–1993. 12

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu (2021): "Notching R&D investment with corporate income tax cuts in China," *American Economic Review*, 111, 2065–2100. 2, 3

Cornelli, Francesca and Mark Schankerman (1999): "Patent renewals and R&D incentives," *The RAND Journal of Economics*, 197–213. 4

Dang, Jianwei and Kazuyuki Motohashi (2015): "Patent statistics: A good indicator for innovation in China? Patent subsidy program impacts on patent quality," *China Economic Review*, 35, 137–155. 2, 3

Economist, The (2020): "Chinese inventiveness shows the weakness of the law," *The Economist*. 8

Fitzgerald, Tristan, Benjamin Balsmeier, Lee Fleming, and Gustavo Manso (2021): "Innovation search strategy and predictable returns," *Management science*, 67, 1109–1137. 26

Ganco, Martin, Rosemarie H Ziedonis, and Rajshree Agarwal (2015): "More stars stay, but the brightest ones still leave: Job hopping in the shadow of patent enforcement," *Strategic Management Journal*, 36, 659–685. 4

Gilbert, Richard and Carl Shapiro (1990): "Optimal patent length and breadth," *The RAND Journal of Economics*, 106–112. 4

Grossman, Gene M and Elhanan Helpman (1991): "Quality ladders in the theory of growth," *The review of economic studies*, 58, 43–61. 4

Hall, Bronwyn H and Rosemarie H Ziedonis (2001): "The determinants of patenting in the US semiconductor industry, 1980-1994," *Rand Journal of Economics*, 32, 101–128. 4

Helmers, Christian et al. (2018): *The economic analysis of patent litigation data*, vol. 48, WIPO. 8

Hopenhayn, Hugo A and Matthew F Mitchell (2001): "Innovation variety and patent breadth," *RAND Journal of Economics*, 152–166. 4

Hu, Albert (2014): "Propensity to patent, competition and China's foreign patenting surge," in *Intellectual Property for Economic Development*, Edward Elgar Publishing. 3

Hu, Albert Guangzhou and Gary H Jefferson (2009): "A great wall of patents: What is behind China's recent patent explosion?" *Journal of Development Economics*, 90, 57–68. 3

Hu, Wei, Tohru Yoshioka-Kobayashi, and Toshiya Watanabe (2020): "Determinants of patent infringement awards in the US, Japan, and China: A comparative analysis," *World Patent Information*, 60, 101947. 8

Jensen, Richard and Marie Thursby (2001): "Proofs and prototypes for sale: The licensing of university inventions," *American Economic Review*, 91, 240–259. 26

Jones, Charles I. (1995): "R & D-Based Models of Economic Growth," *Journal of Political Economy*, 103, 759–784. 16

Klemperer, Paul (1990): "How broad should the scope of patent protection be?" *The RAND Journal of Economics*, 113–130. 4

Lamin, Anna and Miguel A Ramos (2016): "R&D investment dynamics in agglomerations under weak appropriability regimes: Evidence from Indian R&D labs," *Strategic management journal*, 37, 604–621. 4

Lerner, Josh (2009): "The empirical impact of intellectual property rights on innovation: Puzzles and clues," *American Economic Review*, 99, 343–48. 4

Li, Xibao (2012): "Behind the recent surge of Chinese patenting: An institutional view," *Research policy*, 41, 236–249. 2, 3

March, James G (1991): "Exploration and exploitation in organizational learning," *Organization science*, 2, 71–87. 26

Moser, Petra (2005): "How do patent laws influence innovation? Evidence from nineteenth-century world's fairs," *American economic review*, 95, 1214–1236. 4

Nordhaus, William D (1969): "An economic theory of technological change," *The American Economic Review*, 59, 18–28. 4

O'Donoghue, Ted, Suzanne Scotchmer, and Jacques-François Thisse (1998): "Patent breadth, patent life, and the pace of technological progress," *Journal of Economics & Management Strategy*, 7, 1–32. 4, 12

Paik, Yongwook and Feng Zhu (2016): "The impact of patent wars on firm strategy: Evidence from the global smartphone industry," *Organization Science*, 27, 1397–1416. 4

Qian, Yi (2007): "Do national patent laws stimulate domestic innovation in a global patenting environment? A cross-country analysis of pharmaceutical patent protection, 1978–2002," *The Review of Economics and Statistics*, 89, 436–453. 4

Rozelle, Scott and Natalie Hell (2020): "Invisible China," in *Invisible China*, University of Chicago Press. 22

Sun, Zhen, Zhen Lei, Brian D Wright, Mark Cohen, and Taoxiong Liu (2021): "Government targets, end-of-year patenting rush and innovative performance in China," . 2, 3

Thursby, Jerry G and Marie C Thursby (2002): "Who is selling the ivory tower? Sources of growth in university licensing," *Management science*, 48, 90–104. 26

Wei, Shang-Jin, Jianhuan Xu, Ge Yin, and Xiaobo Zhang (2021): "Bureaucratic Bean Counting and Patent Subsidies: A Welfare Evaluation of China's Pro-innovation Program," . 2, 3

Yamaguchi, Kazuhiro (2010): "Japanese Patent Litigation and Its Related Statistics—Current Environment and Future Agenda—," *Notes*, 2011, 2013. 8

Zhao, Minyuan (2006): "Conducting R&D in countries with weak intellectual property rights protection," *Management science*, 52, 1185–1199. 4

A Tables

Table 5: Calibration of Inventive Steps for Incremental Versus Radical Innovation

	U.S. family patents of Chinese patents	U.S. patents	U.S. patents (self citations excluded)	Chinese patents with U.S. family patents	Chinese patents without U.S. family patents
1%	1	1	1	1	1
25%	2	3	2	2	2
50%	5	6	5	4	3
75%	11	15	14	7	6
95%	34	61	56	18	14
99%	88	170	159	30	25
Mean	10.05	16.81	15.4	5.68	4.8
Std	18.13	48.63	45.59	6.69	5.42
N	18068	1133267	1094413	19563	158560

Table 6: Distribution of Forward Citations Received by Invention Patents Made by Domestic Chinese Firms

Application year	Obs	Mean	Std. dev.	Min	Max
2001	6,398	3.718975	4.505303	1	146
2002	11,210	4.103747	4.940329	1	121
2003	16,097	4.194819	5.408678	1	370
2004	20,398	4.215953	4.639633	1	85
2005	28,206	4.361306	4.668211	1	147
2006	39,861	5.155214	5.906086	1	240
2007	56,241	5.770506	6.215057	1	249
2008	63,718	8.050159	9.300262	1	501
2009	63,915	7.441649	8.326586	1	220
2010	55,577	6.205553	8.428729	1	647
2011	79,287	5.562501	7.285456	1	536
2012	135,086	5.594577	6.80842	1	535
2013	210,072	5.314545	6.301471	1	258

Note: *Forward citations received by patents filed before 2007 (including) are counted within 10 years since the patent application years. For the rest patents, forward citations are the total number of citations the patent has received till the data collection time (year 2021).*

Table 7: Patent Production Function Estimates for R&D Manufacturing Firms (2005-2007, 1000RMB)

Form of variables	Standardized		Log(x+1)		Raw number	
DV	family patents	non-family patents	family patents	non-family patents	family patents	non-family patents
	(1)	(2)	(3)	(4)	(5)	(6)
R&D expenditure	0.757***	0.927***	0.0507***	0.161***	0.000196***	0.000384***
	(0.253)	(0.262)	(0.0101)	(0.0134)	(5.39e-05)	(0.000107)
Per-patent R&D expenditure	**5503.09**	**2594.63**			**5102.04**	**2604.17**
#employees	0.275	-0.112	0.0267**	0.112***	-4.770	-11.64**
	(0.379)	(0.101)	(0.0133)	(0.0170)	(4.023)	(5.039)
assets/debts	-0.0550	-0.0198	0.000114	0.0432	-6.120**	-8.841**
	(0.0749)	(0.0234)	(0.0220)	(0.0371)	(2.864)	(3.932)
ROA	-0.575	-1.561	0.0588	-0.0964	1.901	-5.152
	(3.103)	(1.895)	(0.0554)	(0.0829)	(5.649)	(7.365)
Constant	-6.000	-0.701	-0.825***	-2.276***	-39.73	-52.56
	(5.560)	(1.292)	(0.224)	(0.276)	(49.40)	(59.81)
2-digit industry FEs	Y	Y	Y	Y	Y	Y
Observations	2,403	2,403	2,402	2,402	2,402	2,402
R-squared	0.561	0.680	0.096	0.222	0.541	0.682

Note: *Standardized estimates report the beta coefficients. Log(x+1) estimates report results using logarithm transformation of all variables. Non-zero sample incorporates firms with positive R&D expenditures as well as positive number of patents. Patents refer to invention patents only, counted by their filing years. 1-SD for R&D expenditures, family patents and non-family patents are 208291.9 (1k RMB), 50, 86.6, respectively. Robust standard errors are shown in parentheses:* *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Proofs

B.1 Enforcement Damages

If an innovator arises with step size $\gamma < B$ and decides to produce nonetheless, then they will displace the incumbent firm and earn normalized profits $\tilde{\pi} = 1 - \gamma^{-1}$. Absent legal intervention, they can expect to earn these profits until someone introduces a better technology. Since they will not have patented, it is reasonable to suppose that subsequent innovators will continue to build off of the original patent. Thus they will be displaced when someone generates an innovation with step size larger than γ, which will happen at the modified rate $\eta_k R_k^{1-\beta} \gamma^{-\alpha_k}$, leading to a present value of production at any time of

$$\frac{1 - \gamma^{-1}}{\rho + \eta_k R_k^{1-\beta} \gamma^{-\alpha_k}} = \frac{1 - \gamma^{-1}}{\rho + \tau(\gamma/B)^{-\alpha_k}}$$

which is strictly increasing in γ. Thus the largest expected value will arise for the largest value of γ, which is $\gamma = B$ in this case. Thus the maximal expected value will be

$$\tilde{D}(B) = \frac{1 - B^{-1}}{\rho + \tau}$$

Note that because of the normalization we are using, this is implicity considered as a fraction of average yearly product line revenue (since our discount rate is yearly). Thus if we wished to enforce a breadth of $B = 5\%$, given $\rho = 5\%$ and $\tau = 15\%$, we would need to set damages roughly equal to 25% of a firm's yearly revenue for a single product.

B.2 Patent Fees

The resulting type-conditional expected gain from innovation is now

$$\tilde{z}_k = \int_B^\infty \left[\tilde{v}_k(\gamma) - \tilde{d}\right] dF_k(\gamma) = \frac{B^{-1}}{1 + \alpha_k} \frac{B^{-\alpha_k}}{\rho + \tau}$$

The additional B^{-1} factor relative to the baseline arises because firms do not face competition from the partially imitating prior incumbent. This will nonetheless be cancelled out by wage effects in the free entry condition.

Because the effective B value depends on τ, we will actually see a smooth transition between incremental and radical innovation resulting from a sort of mixed strategy. Thus we need to entertain a slightly more general free entry condition, letting $R = R_i + R_r$. Said condition is still derived from $R^{-\beta}\eta_k\tilde{z}_k = (1-s)\tilde{w}$, meaning

$$\frac{\eta_k B^{-1}}{1+\alpha_k}\frac{R^{-\beta}B^{-\alpha_k}}{\rho+\tau} = \frac{(1-s)\Lambda^{-1}}{1-R} \tag{4}$$

For sufficiently large or small $\tilde{d}$ either incremental or radical innovation will prevail. In this case, we can show that the equation characterizing the equilibrium is identical to the B policy case. Because of the aforementioned markup effects, the distribution of λ will shifted to the right by a factor B so that $\Lambda^{-1} = B^{-1}\frac{\alpha_k}{1+\alpha_k}$, and thus we find

$$\frac{\eta_k R_k^{-\beta}B^{-\alpha_k}}{\rho+\eta_k R_k^{1-\beta}B^{-\alpha_k}} = \frac{(1-s)\alpha_k}{1-R_k}$$

and B is implicitly now a function of $\tilde{d}$ and and R_k via $\tau_k = \eta_k R_k^{1-\beta}B^{-\alpha_k}$. Letting τ_i^* and τ_r^* be the innovation rates for the case where $B = B^*$, these pure regimes will arise when

$$\tilde{d} \leq \tilde{d}_0 \equiv \frac{1-1/B^*}{\rho+\tau_i^*} \quad \text{or} \quad \tilde{d} \geq \tilde{d}_1 \equiv \frac{1-1/B^*}{\rho+\tau_r^*}$$

Conversely, for $\tilde{d} \in (\tilde{d}_0, \tilde{d}_1)$, the effective B will remain constant at B^*, meaning

$$\tilde{\tau} = \frac{1-1/B^*}{\tilde{d}} - \rho$$

and there will be an interior outcome, with some firms engaging in incremental innovation and others engaging in radical innovation. Let H^* denote the common value of $H_k = \eta_k\tilde{v}_k$ at the critical patent breadth B^* or

$$H^* \equiv \left(\frac{\eta_i}{1+\alpha_i}\right)^{\frac{\alpha_r}{\alpha_i-\alpha_r}}\left(\frac{\eta_r}{1+\alpha_r}\right)^{\frac{\alpha_i}{\alpha_i-\alpha_r}}$$

Plugging this into Equation (4), the free entry condition for both types becomes

$$\frac{H^* R^{-\beta}}{\rho + \tilde{\tau}} = \frac{(1-s)B^* \Lambda^{-1}}{1-R} \tag{5}$$

In this mixed setting, letting $f_k = R_k/R$, we can write the aggregate innovation rate as

$$\begin{aligned}
\tilde{\tau} = \tilde{\tau}_i + \tilde{\tau}_r &= \eta_i R^{-\beta} R_i (B^*)^{-\alpha_i} + \eta_r R^{-\beta} R_r (B^*)^{-\alpha_r} \\
&= H^* R^{1-\beta} \left[(1+\alpha_i) f_i + (1+\alpha_r) f_r\right] \\
&= H^* R^{1-\beta} (1+\bar{\alpha})
\end{aligned}$$

where $\bar{\alpha} \equiv \alpha_i f_i + \alpha_r f_r$. With this we can find the average markup, noting that this derivation relies critically on $H_k = H^*$

$$\begin{aligned}
\Lambda^{-1} &= \left(\frac{\tilde{\tau}_i}{\tilde{\tau}}\right)(B^*)^{-1}\left(\frac{\alpha_i}{1+\alpha_i}\right) + \left(\frac{\tilde{\tau}_r}{\tilde{\tau}}\right)(B^*)^{-1}\left(\frac{\alpha_r}{1+\alpha_r}\right) \\
\Rightarrow \quad (B^*)\Lambda^{-1} &= \frac{\alpha_i f_i + \alpha_r f_r}{(1+\alpha_i) f_i + (1+\alpha_r) f_r} = \frac{\bar{\alpha}}{1+\bar{\alpha}}
\end{aligned}$$

Plugging this into Equation (5), we ultimately find an equation characterizing $\bar{\alpha}$

$$\begin{aligned}
\frac{\tilde{\tau}}{\rho + \tilde{\tau}} &= \bar{\alpha}(1-s)\left(\frac{R}{1-R}\right) \\
\Rightarrow \quad R &= \frac{\tilde{\tau}}{\tilde{\tau} + \bar{\alpha}(1-s)(\rho + \tilde{\tau})} \\
\Rightarrow \quad \tilde{\tau} &= H^*(1+\bar{\alpha})\left[\frac{\tilde{\tau}}{\tilde{\tau} + \bar{\alpha}(1-s)(\rho + \tilde{\tau})}\right]^{1-\beta}
\end{aligned}$$

It can be show that the above equation is increasing in $\bar{\alpha}$ if and only if

$$\beta(1+\bar{\alpha}) - s\bar{\alpha} > \frac{\rho}{\rho + \tilde{\tau}}$$

A sufficient condition for this to hold for any $\tilde{\tau}$ is then $\beta(1+\alpha_k) > 1 + s\alpha_k$ for $k \in \{i, r\}$. Given that we are assuming $\alpha_r < \bar{\alpha} < \alpha_i$, in the case where $s < \beta$, it suffices to assume that $(\beta - s)(1+\alpha_r) > 1$.

Given a value for $\bar{\alpha}$, one can derive the implied weights $1 - f_i = f_r = (\bar{\alpha} - \alpha_i)/(\alpha_r - \alpha_i)$.

Finally, we can compute the aggregate growth rate

$$
\begin{aligned}
g &= \left[\frac{1}{\alpha_i} + \log(B^*)\right]\tilde{\tau}_i + \left[\frac{1}{\alpha_r} + \log(B^*)\right]\tilde{\tau}_r \\
&= H^* R^{1-\beta}\left[\left(\frac{1}{\alpha_i} + \log(B^*)\right) f_i + \left(\frac{1}{\alpha_r} + \log(B^*)\right) f_r\right] \\
&= H^* R^{1-\beta}\left[(1+\bar{\alpha})\log(B^*) + \frac{1+\hat{\alpha}}{\hat{\alpha}}\right]
\end{aligned}
$$

where $\hat{\alpha}$ is the harmonic mean of α, defined by $\hat{\alpha}^{-1} = f_i\alpha_i^{-1} + f_r\alpha_r^{-1}$.

B.3 Citation Distribution

To model citations, we specify a distribution for citation counts c conditional on a value for γ. We then compute the unconditional mean and variance of citations under these assumptions. First, the Law of Total Variance implies

$$\mathbb{V}[c] = \mathbb{E}\left[\mathbb{V}[c|\gamma]\right] + \mathbb{V}[\mathbb{E}[c|\gamma]]$$

We will also need the following properties of the Pareto distribution

$$\mathbb{E}[\gamma - 1] = \frac{1}{\alpha - 1} \quad \text{and} \quad \mathbb{V}[\gamma - 1] = \mathbb{V}[\gamma] = \frac{\alpha}{(\alpha-1)^2(\alpha-2)}$$

Our particular assumption for the citation data generating process is that both conditional mean and variance are linear in the step size, so that $\mathbb{E}[c|\gamma] = \mathbb{V}[c|\gamma] = \kappa(\gamma - 1)$, which would be consistent with a Poisson distribution. The unconditional mean of citations is then straightforward to calculate

$$\mathbb{E}[c] = \mathbb{E}[\mathbb{E}[c|\gamma]] = \mathbb{E}[\kappa(\gamma-1)] = \frac{\kappa}{\alpha - 1}$$

For the unconditional variance, we use the equation above

$$\mathbb{V}[c] = \mathbb{E}[\kappa(\gamma-1)] + \mathbb{V}[\kappa(\gamma-1)] = \frac{\kappa}{\alpha-1} + \frac{\kappa^2\alpha}{(\alpha-1)^2(\alpha-2)}$$

With this we can compute the modified index of dispersion that depends only α but not κ

$$\tilde{D} = \frac{\mathbb{V}[c] - \mathbb{E}[c]}{\mathbb{E}[c]^2} = \frac{\alpha}{\alpha - 2}$$

Note that existing indices of dispersion, such as $\mathbb{V}[c]/\mathbb{E}[c]$ or $\sqrt{\mathbb{V}[c]}/\mathbb{E}[c]$ depend on κ, which is a nuisance parameter in our case. Thus we use the modified index of dispersion ($\tilde{D}$) for the calibration.