Unleashing the Dragon: The Case for Patent Reform in China

Lee G. Branstetter* Douglas Hanley† Huiyan Zhang‡

September 27, 2022

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. Based on these theoretical predictions, we make two policy suggestions to further accelerate China's economic and technological progress. First, the distribution of patent damage awards should be significantly increased. Second, the quality of patented inventions could be improved by increasing patent application fees.

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.[1]. Despite the intensity and longevity of these

*Carnegie Mellon University

†University of Pittsburgh

‡Carnegie Mellon University; Catolica Lisbon School of Business and Economics

[1] See Economist (2020) for one example of domestic criticism of the system. The Chinese patent office conducts annual surveys of inventors of Chinese patents which detail many complaints. Full

complaints and concerns, the debate over China's patent system has long suffered from a lack of focus on exactly where and how it falls short. Though a recent wave of (mostly empirical) research has made substantial progress in terms of documenting specific shortcomings of Chinese patent policy (Z. Chen et al., 2021; Dang and Motohashi, 2015; Li, 2012; Sun et al., 2021; Wei et al., 2021), the literature still lacks theoretical and empirical rigor in terms of analyzing 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 the exploration on the consequences for innovation, economic growth and social welfare.

We start with a description of China's patent system that documents some important new 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. While this may be true of patent systems in general, it is especially true for Chinese patents. 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. If the major benefit conferred upon a patent holder is the ability to credibly threaten to punish infringers to a degree that deters infringement, then this benefit is all but absent in China. 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

reports are all in Chinese, available in https://www.cnipa.gov.cn/col/col88/index.html growth-theoretic model in which Chinese patent policy determines the nature and extent of firm innovative efforts. In this model, the weaknesses of the Chinese patent system identified in Section 2 interact with one another, resulting in insufficiently broad patent protection. The model shows how this distorts innovative effort. 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. The low infringement awards are insufficient to deter infringement of high quality inventions. This certain 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. Using a comprehensive dataset containing all Chinese invention patents filed by domestic Chinese firms from 2001 to 2007, 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. Although we do not explicitly consider this in the current closed economy version of our model, slower technological progress in China almost certainly lowers global welfare and the absence of effective patent enforcement and the ubiquity of undeterred infringement in China plausibly leads to persistent IP frictions with Chinese trading partners that further damage Chinese and global welfare.

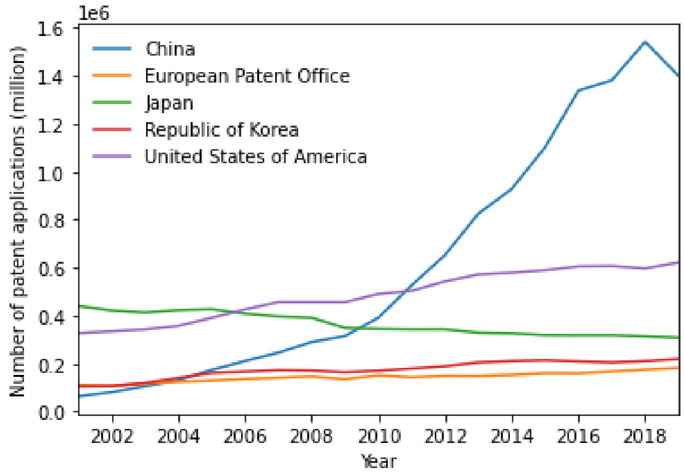
We then consider the implications of our model for future Chinese patent policy. We argue that the effective breadth or scope of protection for substantive or radical inventions must be increased. How can this be achieved? First, the distribution of patent damage awards must shift significantly to the 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 (Z. Chen et al., 2021; Dang and Motohashi, 2015; A. Hu, 2014; A. G. Hu and Jefferson, 2009; Li, 2012; Sun et al., 2021; Wei et al., 2021). 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 (Branstetter and Sakakibara, 2002; Hall and Ziedonis, 2001; Lerner, 2009; Moser, 2005; Qian, 2007). Prior research has explored optimal patent length and breadth (Gilbert and Shapiro, 1990; Klemperer, 1990; Nordhaus, 1969). Our research has been particularly influenced by the work of O'Donoghue et al. (1998), who entertain 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 (Belderbos et al., 2021; Branstetter et al., 2006; Lamin and Ramos, 2016). As a result, innovating firms may need alternative innovation strategies to protect their inventions (Beukel and Zhao, 2018; Paik and Zhu, 2016; Zhao, 2006). 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 firms own products and external innovation to capture new

Figure 1: Total patent applications by filing office



Source: WIPO statistics database. Last updated: January 2021.

product lines; Acemoglu et al. (2018) consider a setting in which firms vary in their intrinsic aptitude for generating new 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.

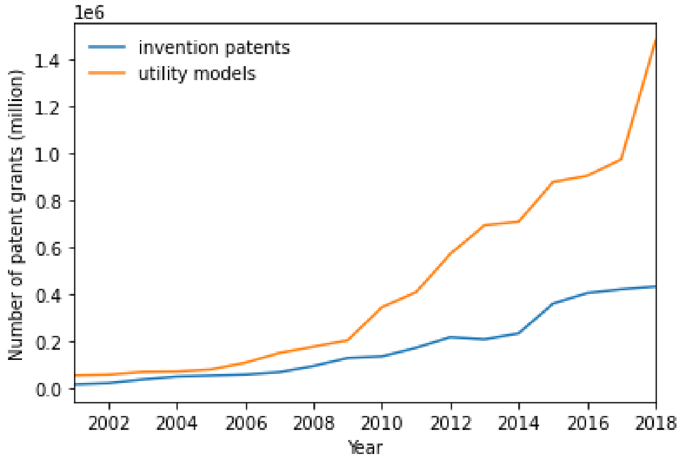
2 Key Features of China's Patent System

2.1 A strong bias toward incremental innovation

China did not establish a modern patent system until 1985. However, it has witnessed incredibly rapid growth in patent applications since then. Figure 1 shows the recent trend of patent applications received by the top five patent offices in the world. It is evident that China's growth in patenting accelerated around 2010 and two years later China bypassed the U.S., ranking first in the world in terms of the number of both patent applications and patent grants ever since.

However, a variety of other indicators suggest that the quality of patents filed in China

Figure 2: Total patent grants of inventions and utility models by China's patent office

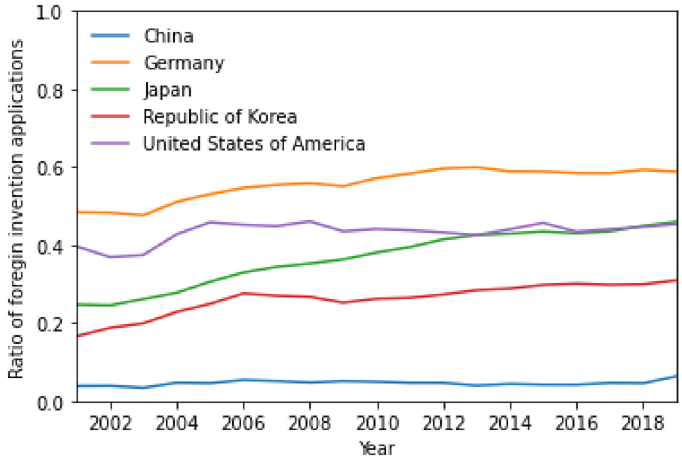


Source: China's Patent Office.

are far less impressive than their quantity. Figure 2 shows the evolution paths of two types of patents granted by China's patent office. 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. Most patent systems, including that of the U.S., do not even grant patent protection of any kind to these utility model "innovations."

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 3 shows that inventors in most countries decide to forego foreign patent protection for a large fraction of their domestically patented inventions. This partly reflects the incremental nature of much patented invention everywhere and the degree to which an additional year gives inventors a better signal of the true economic value of their inventions, something not always known with certainty at the time of initial filing. However, China stands out in comparison to other major innovating countries. German inventors seek foreign protection for about 60% of their inventions. American and Japanese inventors seek foreign protection for slightly less than half of their inventions. Korean inventors seek foreign protection for about 30% of their inventions, a fraction that has risen significantly over time. In striking contrast,

Figure 3: The ratio of invention patents filed abroad over all patents filed by applicant origins



Source: WIPO statistics database. Last updated: January 2021.

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, and has been for years, having long ago surpassed Germany, the United States, and Japan in terms of its total sales in foreign markets. 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. This would appear to be a striking vote of "no confidence" in the true novelty and value of these inventions by those parties best positioned to make such a judgment.

Apart from invention patents, the common patent type adopted by most countries, utility model patents in China show a similar pattern in that the recent increase in applications is dominated by only domestic filings. In contrast, both Japan and Germany, two other major patent systems that grant utility models, have seen an increasing tendency for their utility model applications to be protected abroad as well as at home. Despite this much higher tendency to protect utility models abroad as well as at home, Japanese and German inventors register far fewer utility models relative to invention patents than do Chinese inventors, and the relative popularity of utility models has declined in Japan and Germany, even as it has surged in China. In recent years, Japanese inventors have only filed about 2% as many utility models as invention patents, and their German counterparts now file fewer than 6% as many utility models as invention patents (WIPO statistics database [2]).

All of this supports our view that contemporary Chinese patented invention is overwhelmingly incremental in nature. Anecdotal evidence of the incremental nature of much of Chinese invention abounds beyond the boundaries of China's patent system. China-based factories manufacture most of the world's IT hardware, but much of it is based on foreign designs (Lovely, Liang, et al. (2018)). China-based digital giants with business models closely resembling those of foreign pioneers dominate their vast internal market, but, with a few exceptions, are all but absent outside the confines of the Chinese internet (Dvorak and Woo, 2019). Chinese firms run a huge surplus in manufactured goods but a huge deficit in intellectual property licensing fees (Kennedy, 2017). The conspicuous over-representation of inventors and scientists of Chinese descent in pathbreaking inventions and fundamental scientific advances outside China strongly suggests that what exists in China is a response to the incentives inventors face rather than any fundamental constraint on Chinese inventive creativity. We will return to this point later in the paper.

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 (USD 1,553 to 155,300) by the patent reform implemented in 2008 [3], in practice, the average compensation received by invention patent owners who seek infringement remedies is only around 440,000 RMB (USD 68,000).[4].

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

[2]See https://www.wipo.int/ipstats/en/

[3]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.

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

Table 1: Summary statistics of China's patent infringement lawsuits

	Inventions	Utility models	Designs	overall
Average damage awards (2021 US$)	68,000	16,838	7,606	17,400
Average proceeding time (months)	8.6	6.2	5.1	5.9
% compensation received/compensation claimed	38%	28%	29%	33%
Win rate	63%	71%	83%	77%

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

Table 2: Comparison of patent litigation statistics in China, Japan, and the U.S.

	China	Japan	US
Damage awards (2010 US$)			
min	1291	44	1873
25th percentile	15838	27201	500000
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
Average proceeding time of 1st instance (months)	18-Jun	15-Dec	18-42
Average lawsuit costs (US $)	20,000-150,000	300,000-500,000	1,000,000-6,000,000
Average damage awards/ average 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 (W. Hu et al., 2020). The rest of the figures are from a WIPO report (Helmers et al., 2018).

firm. An 80 square meter apartment in Beijing's Xicheng district could go for USD 1.5 million. As it is, 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.

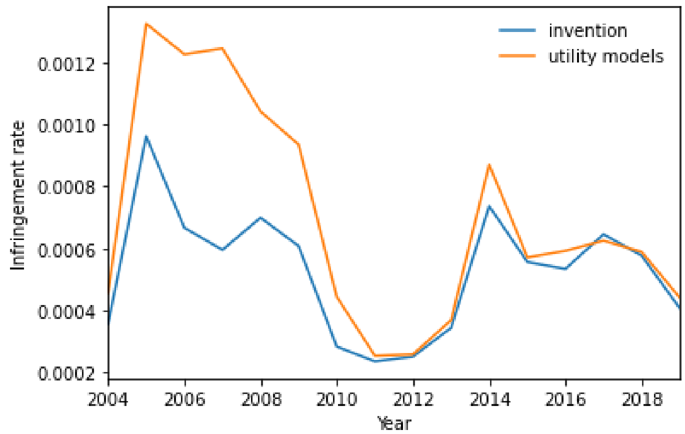
A comparison with other major patent systems demonstrates that China's system is characterized by extremely low damage awards. Table 2 shows a comparison of China, Japan and the US in terms of major features of patent litigation practice. At every point of the distribution, damage awards in China are 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. levels. The patent system adopted by China in the 1980s was based on Japan's system at the time, and even the modern-day 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 ("Chinese inventiveness shows the weakness of the law", 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 amount of 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 ratios of Japan and US are all above one, suggesting that the expected returns received by 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 the 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 puny 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 just 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. Despite well-publicized growth in the number of actual patent lawsuits, contemporary China is characterized by an unprecedented and remarkably pronounced *decreasing* trend in the patent litigation rate- the ratio of patent litigation case numbers in year t divided by the number of patents in force in that year. Both inventions and utility

Figure 4: Patent infringement rate in China



Note: Patent infringement rate is calculated by dividing the number of patent infringement cases filed in year t over the total number of patents in force in year t. Data source: IncoPat data, https://www.incopat.com/

model patent litigation cases have increased in recent years, but when we divide the infringement case numbers by the count of patents in force each year, shown as the litigation rate in figure 4, we see a clear decline in recent years. The magnitude of this infringement rate, even at its peak-around 0.13 percent, is much lower than the corresponding level in other countries. By comparison, the litigation rate in US is reported to be around 4 percent (Bessen and Meurer, 2005). The low and rapidly declining infringement rate in China is understandable if infringement damages are too low to deter infringement ("Chinese inventiveness shows the weakness of the law", 2020).

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, as has already been noted. Japan's utility model grants amount to only 2% of its invention patent grants, and Germany's amount to less than 6% of its invention patent grants. 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 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 compared with their American and European counterparts, and this gap is increasing over time. By 2016, Chinese examiners were evalutaing 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 decisions on granting or rejecting a patent application given these burdens and the relatively low level of monthly salary Chinese patent examiners receive (anecdotal evidence suggests a salary around 10,000 RMB, which is USD 1,600) by international standard. 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 exaggerate the poor 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 until very recently to encourage patent applications. China's central and local governments have implemented a variety of patent subsidies and R&D tax credit policies to encourage patent application regardless of quality of the incentivized applications (Li, 2012; Dang and Motohashi, 2015; Z. Chen et al., 2021; Wei et al., 2021; also see Appendix B for details). These include a substantial reduction in the corporate income tax rate offered up by the central government, direct subsidies for patent applications offered up by local governments at various levels, preferential access to a market listing at home and/or abroad, and potentially large effects on the likelihood of getting financed by state-owned or state-directed financial institutions, including the largest commercial banks in China. Added together, these subsidies and near-subsidies create a strong financial case for increased (domestic) patenting activity, but the incentive to patent is significantly disconnected from the motivation of legally protecting substantial innovations from imitation and from incentives to aim for more substantive and fundamental rather than incremental innovations.

In the very recent past, the Chinese central government has taken the important, positive step of prohibiting direct subsidies of patent applications by any level of government. Unfortunately, this directive has not been fully implemented by local governments, and the indirect subsidies embedded in Chinese corporate tax law continue to create strong incentives for patent applications of sufficient number to overwhelm the evaluation resources of the Chinese patent office.

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. In pursuit of analytical simplicity and clarity, 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. Since all technological progress can be represented in these models as movement along a line (the eponymous "quality ladder") that measures the technology of the leading producer at every point in time, we will represent incremental and radical innovation as being differentiated by the distributions of inventive step sizes along this line associated with each of these two modes of innovation. [5] The simple, unidimensional treatment of technological progress in quality-ladder models facilitates the derivation of closed form solutions and yields a strong and intuitive relationship between the effective strength and breadth of patent protection and the nature of technological progress.

We assume that 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 China.) The patent system itself facilitates this knowledge transfer, by requiring inventors to disclose

[5] The use of the term "radical innovation" may to 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 along the quality ladder than the incremental innovations with which we are comparing them.

the workings of their technologies. However, disclosure may not be balanced by the appropriate breadth of protection in a weak patent system. In this context, 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. When patents are nonexistent or weak, this chain of events could dramatically undermine a firm's incentives for radical innovation, especially if it is costlier and/or riskier than incremental innovation. We will show that this, in turn, leads to an equilibrium in which even firms with the capacity to engage in radical innovation nevertheless invest only in incremental innovation.

The introduction of broad, effectively enforced patents restores the incentives for radical innovation. Patents allow innovators to exclude would-be entrants who offer only minor improvements to an innovation. A market leader can still be displaced, but only by an innovator that offers a substantive improvement. With broad and effective patent protection, in equilibrium, the longer expected life of a "radical innovation" in the marketplace compensates radical innovators for the higher costs and greater risks they incur. Intuitively, technological progress becomes characterized by fewer steps, but longer strides along the quality ladder. In our model, as we will show, 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 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.[6]

Similarly to our paper, Y. Chen et al. (2018) focus on the distinction between "safe"

[6] 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.

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 much more risky than incremental. In their model, the safe innovation mode is truly riskless and yields an innovation of a fixed size, while the risky mode yields an innovation of random size. Thus for modestly sized patent breadths (as they assume throughout), the safe mode is unaffected while the risky mode is left-censored, thus decreasing its expected return. Nonetheless, our model does not nest theirs, as they allow for an arbitrary bounded distribution of risky innovation outcomes, rather than Pareto, as we assume here.

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 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. Thus, the profit level of a given final good aggregator will be

$$\Pi = \exp\left(\int_0^1 \log(y_j)dj\right) - \int_0^1 p_j y_j dj$$

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 a $\sim$ denote 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 $y = PQ$ and that in any case, we must have $y \leq PQ$ due to Jensen's inequality. In this way, the ratio Λ/Ω can be seen as a measure of labor misallocation across product lines.

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

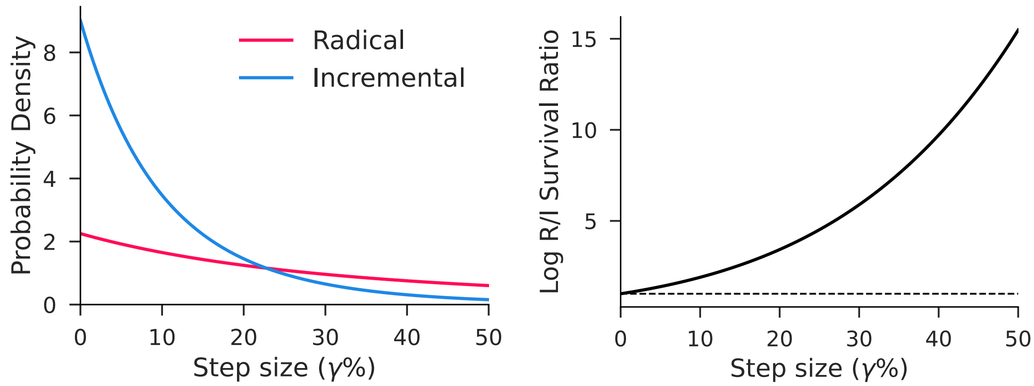
$$F_k(\gamma) = 1 - \gamma^{-\alpha_k} \quad \Rightarrow \quad f_k(\gamma) = \alpha_k \gamma^{-\alpha_k - 1}$$

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 5, 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.

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 the the case that patent infringement damage awards are large enough to deter entry by any incremental innovator or series of incremental innovators whose modification of the focal innovation

Figure 5: Comparison of incremental and radical innovation step size distributions for case where $\alpha_i > \alpha_r$.



places it within B units of that innovation. Without deterrence, minor innovators could quickly displace a "radical" innovator, undermining the rewards for radical innovation. To simplify exposition, we do not explicitly model the game theory of deterrence here. However, in the appendix we present a sketch model demonstrating 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.

Credible defense of intellectual product space B also requires that the patent system be able to recognize when a follow-on innovator's effort lies within B units of the focal innovation versus when it lies beyond. Large damage awards will not bring effective deterrence unless the incumbent can prove that an infringer is actually infringing. To simplify exposition, we do not model the game theory of entry deterrence when agents and the legal system operate with imperfect information. Section 4.1 brings imperfect evaluation of patent boundaries into consideration, making it clear that a certain degree of accuracy in patent evaluation is necessary in order to sustain radical innovation in equilibrium.

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

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

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

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 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.

The government can also subsidize innovation costs. We denote this subsidy rate by $s \in [0,1]$. The costs of this subsidy program are assumed to be funded by a lump-sum tax on consumers that perfectly 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. First, assume that these happen in the aggregate at rate τ_k for $k \in \{i, r\}$. Then the value of a product line is in general

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

where τ is the combined rate of creative destruction from both incremental and radical innovation. Letting $g_v \equiv \frac{\dot{v}_j}{v_j}$, 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 of $r = \rho + g$ then yields

$$v_j = \frac{\pi_j}{\rho + \tau} \quad \Rightarrow \quad \tilde{v}_k = \frac{\tilde{\pi}_j}{\rho + \tau}$$

Thus the value depends only on λ_j via $\tilde{\pi}_j$, meaning we can write the value in a way that is a function of a generic 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. This leads to the value innovation being given by

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

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 $\tau = \tau_i P_i + \tau_r P_r$ 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. 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 $\eta_k \tilde{z}_k$. Thus, substituting in for what we have found so far, 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 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 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. In the case where $\beta = 0$ (no researcher heterogeneity) or in the limit of $\beta \to 1$, one can also find a closed form expression for R_k. For instance, when $\beta = 0$, we arrive at

$$R_k^0 = \frac{1 - (1-s)\left(\frac{\rho\alpha_k}{\eta_k}\right) B^{\alpha_k}}{1 + (1-s)\alpha_k}$$

For general β, 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 amounts to

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

In general, it will be useful to define $H_k \equiv \frac{\eta_k B^{-\alpha_k}}{1+\alpha_k}$ as the payoff relevant term for each innovation type.

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

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

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

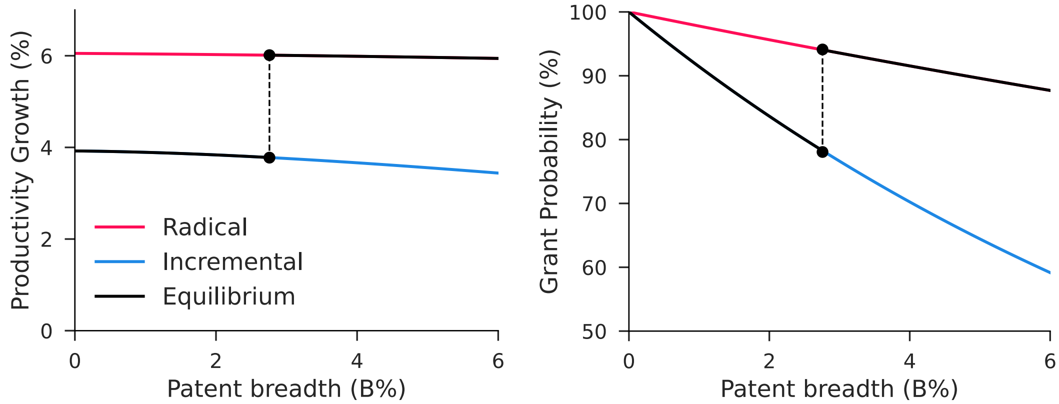
$$B^* = \left[\frac{\eta_i}{\eta_r} \frac{1+\alpha_r}{1+\alpha_i}\right]^{\frac{1}{\alpha_i - \alpha_r}} > 1$$

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. Additionally, the formula for B^* above requires information about the step size distribution (α_k) and the cost level of innovation η_k. In the next section, we will solve the social planner's problem and derive a B policy that requires only knowledge of α_k.

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 $\alpha_i = 4$, $\alpha_r = 3$, $\eta_i = 8$, and $\eta_r = 6$. As can be seen in Figure 6, 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.

Figure 6: Output growth and grant probability as a function of patent breadth policy (B)



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

To understand the net effect of these changes, in Figure 7 we plot the consumption equivalent 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 production arising from changes in production labor levels and the misallocation of production factors due to monopoly distortions. It is clear that there are potentially very large gains associated with a switch to radical innovation. However, only a fraction of these gains are realizable with a patent breadth policy, as in order to induce radical innovation, we must set a relatively large patent breadth, which dampens innovation incentives and leads us to discard smaller innovations.

3.6 Social Planner

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

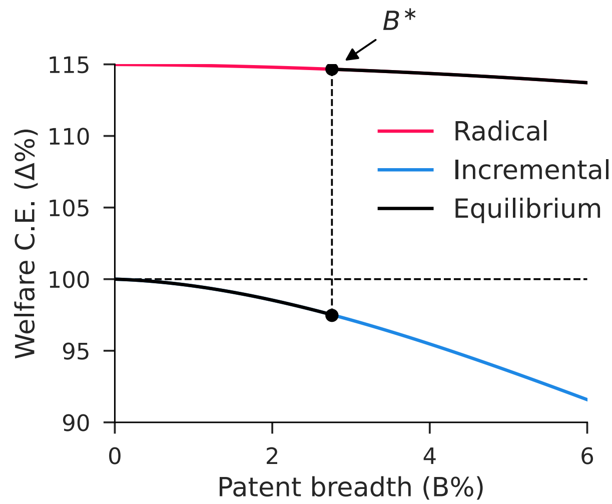
$$g = \tau_i \mathbb{E}_i[\log(\gamma)] + \tau_r \mathbb{E}_r[\log(\gamma)]$$

Given our distributional assumptions, we know that $\mathbb{E}_k[\log(\gamma)] = \frac{1}{\alpha_k}$, and so the growth rate is ultimately

$$g = \frac{\tau_i}{\alpha_i} + \frac{\tau_r}{\alpha_r}$$

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

Figure 7: Aggregate welfare levels as a function of patent breadth policy (B)



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

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

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

Using the typical approach, we find the conditions for optimality of R_i and R_r, as well as an equation describing the evolution of the costate variable μ corresponding to the state variable Q

$$\begin{aligned}
0 = H_{R_i} &= -\frac{1}{P} + \mu Q\left[R^{-\beta}\frac{\eta_i}{\alpha_i} - \beta R^{-\beta-1}\left(\frac{\eta_i R_i}{\alpha_i} + \frac{\eta_r R_r}{\alpha_r}\right)\right] \\
0 = H_{R_r} &= -\frac{1}{P} + \mu Q\left[R^{-\beta}\frac{\eta_r}{\alpha_r} - \beta R^{-\beta-1}\left(\frac{\eta_i R_i}{\alpha_i} + \frac{\eta_r R_r}{\alpha_r}\right)\right] \\
\rho\mu - \dot{\mu} = H_Q &= \frac{1}{Q} + \mu g
\end{aligned}$$

Given the frictionless choice between incremental and radical innovation, we would once again expect a so-called "bang-bang" result, whereby only 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 can simplify these equations to

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

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

$$
\begin{aligned}
& \frac{1}{\mu Q} - (\rho - g) = -\frac{\dot{\mu}}{\mu} = g \quad \Rightarrow \quad \mu Q = \frac{1}{\rho} \\
\Rightarrow \quad & \frac{\eta_k}{\rho}\frac{R_k^{-\beta}}{1-\beta} = \frac{\alpha_k}{1-R_k}
\end{aligned}
$$

Again, we can see that this has a unique and interior solution, and in the case where $\beta = 0$, we get the closed form solution

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

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

3.7 Optimal Policy

The distortion factor of the equilibrium determination, which amounts to a comparison of $\tilde{z}_i$ and $\tilde{z}_r$, relative to the optimum is then

$$
G(\alpha_k, B) \equiv \left(\frac{\alpha_k}{1+\alpha_k}\right) B^{-\alpha_k}
$$

One natural objective would be neutralize this distortion for any values of η_i and η_r by setting B so as to satisfy the equation

$$
\begin{aligned}
& G(\alpha_i, \hat{B}) = G(\alpha_r, \hat{B}) \\
\Rightarrow \quad & \hat{B} = \left[\frac{\alpha_i}{\alpha_r} \cdot \frac{1+\alpha_r}{1+\alpha_i}\right]^{\frac{1}{\alpha_i - \alpha_r}}
\end{aligned}
$$

Note that this always satisfies $\hat{B} \geq 1$ and reaches $\hat{B} = 1$ if and only if $\alpha_i = \alpha_r$.

The above value for B always ensures that the type of innovation seen in the social planner's solution prevails in equilibrium. To see this, note that

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

Thus when radical innovation is optimal, we will have $\hat{B} > B^*$ and hence radical innovation will prevail in equilibrium. Alternatively, when incremental innovation is optimal, we will have $\hat{B} < B^*$, meaning incremental will prevail in equilibrium.

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. Further, setting such a policy will involve "throwing away" innovations of size smaller than B.

That said, one could couple this with a 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) innovation subsidy yields the equilibrium research allocation

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

Thus the 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. This policy will satisfy

$$R_k^*(\hat{s}, \hat{B}) = \hat{R}_k$$

Additionally, since the research subsidy is uniform, it doesn't further distort the margin between incremental and radical innovation, so employing the $\hat{B}$ policy ensures the optimal type of innovation prevails.

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 $\hat{B}$ and a modified $\hat{s}'$ to implement the first best optimum.

4 Model Extensions

In this section, we summarize two extensions to the baseline model. The first 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 entertains an alternative type of patent system in which, rather than relying on patent examiners, simply makes inventors pay fixed fee for patent protection.

4.1 Imperfect Examination

Any assessment of patent novelty requires a judgement on the part of the examiner as to whether it meets the office's standards in that dimension. Given the enormous number of different technological fields and the fact that the ideas being described are (one hopes) at the forefront of those fields, it is reasonable to assume that it is costly to do this at scale and that, in any case, errors will happen.

To model this, we suppose that with probability $1-\varepsilon$ the patent office correctly judges patents to have exceeded the existing breadth B and with probability ε the patent office simply grants the new patent with (potentially type-conditional) probability q_k. Thus the parameter ε is a measure of the accuracy of patent judgements, though it is not the true error rate, as even random evaluations will be correct sometimes.

A full a derivation of the resulting equilibrium is given in Section B.2, and we summarize the results here. First, the new factor ε will enter directly into the expression for the probability of being granted patent protection

$$P_k = \varepsilon q + (1-\varepsilon)B^{-\alpha_k}$$

The generic free entry condition in Equation (2) still applies. However, the new probability of patent granting (P_k) is given above, and the new process for determining which innovations make it to production results in slightly different average markups, summarized by Λ and Ω.

One interesting special case to consider is that when the value of q is set exactly equal to the ex ante type-conditional probability of exceeding the threshold B, namely $B^{-\alpha_k}$. In this case, the probability of acceptance will be unchanged from the complete information specification, but the quality of granted patents will still be lower with higher ε. Importantly, there will also be no additional type-differential distortions induced, so the switching threshold B^* will be the same as that seen in the baseline specification. The resulting aggregate growth rate will be

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

which differs from the baseline growth rate in Equation (3) only in the $1-\varepsilon$ term, reflecting the reduced efficiency of innovation filtering resulting from noisy assessment. Relative to these type-conditional q values, using a single interior value for q will effectively favor incremental innovation and lead to a higher switching point B^*.

4.2 Patent Fees

Now consider the case in which 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. There will be no other lower bound on innovation size imposed by the patent office.

Because total output is growing continually, we must index d to output to achieve stationarity. This 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 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 Section B.3, but the main takeaway is that, contrary to the baseline case, this type of policy will induce 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 $\tilde{d}$.

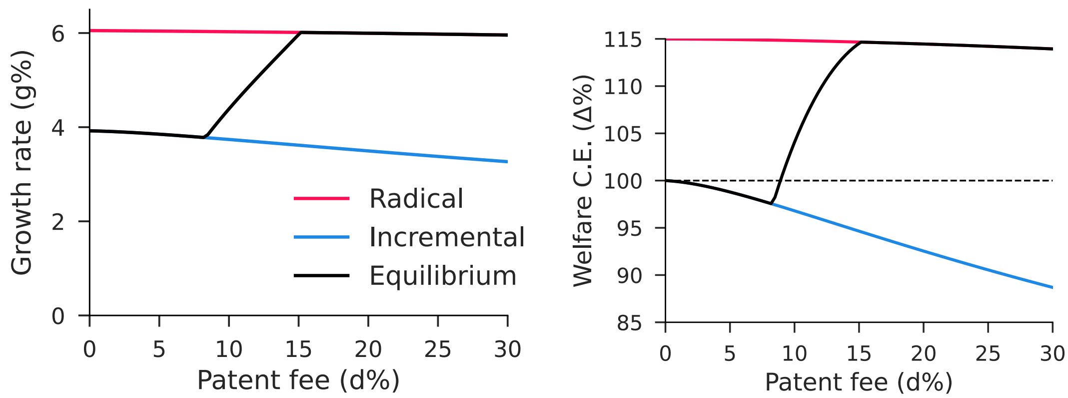


Figure 8: Growth and welfare under various patent fee regimes

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

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

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.

4.3 Additional Extensions

In the Appendix, we cover two additional extensions. The first (Appendix B.4) considers the case where innovation costs (η_k) are not fixed but are instead drawn from some joint distribution at the firm level. This yields outcomes that are continuous in the patent breadth, as opposed to the abrupt transition from incremental to radical that we see in the baseline model. The second (Appendix B.5) replaces incremental innovation with a deterministic outcome so as to isolate the risk channel, and this yields 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 2007. 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.

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., 2019; Azoulay et al., 2010) 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

[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.

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 parameters 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 (Fitzgerald et al., 2021; March, 1991). 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 $\kappa > 0$

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

which would, for instance, 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 Section B.6. Using these, we can construct an index of dispersion that depends only on α but not κ

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

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

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 identifying this parameter from citation data in the case that $\alpha > 2$.

Innovation Cost Parameters. Innovation cost in our model corresponds to a values 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 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.

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

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 A1 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, row 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

[8]Only patents with positive citations are included.

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 exploits 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 of 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 patent records, we regress the proxies for radical and incremental innovations on firms' R&D spending. We relegate the details to Appendix E. Table A2 reports the results. As a robustness check, we try out different forms of 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- over non-family patents- a proxy for incremental innovations- is around 2.12.

One might concern that the data we use for this approach might only reflect the situation more than 10 years ago. Therefore, we try out 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 consequent years. Aside from regressing the patent output on R&D spendings like what we have done in Table A2, here we regress the 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 over 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 the radical innovation could be highly expensive relatively to in-

[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.

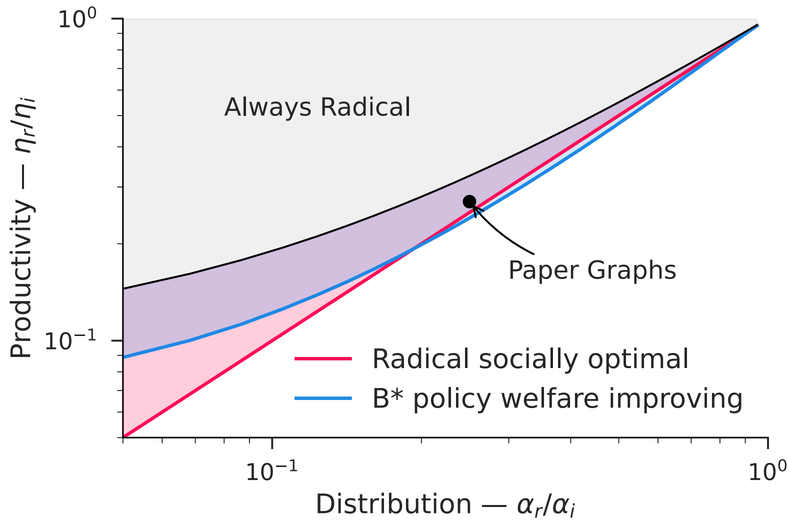


Figure 9: Sensitivity of breadth policy outcome to radical parameters (relatie values in logarithmic scale)

cremental innovation. The gist here is that we do not stick to any specific figure of the parameter calibration and simply explore possible scenarios from 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. These exercises facilitate our discussion on the system under various conditions.

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 9, we plot the relative values of α_r and η_i in logarithmic space. Shaded in gray, there is a region of relatively large η_r values where radical innovation is always the equilibrium outcome, even when $B = 1$, 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 in-

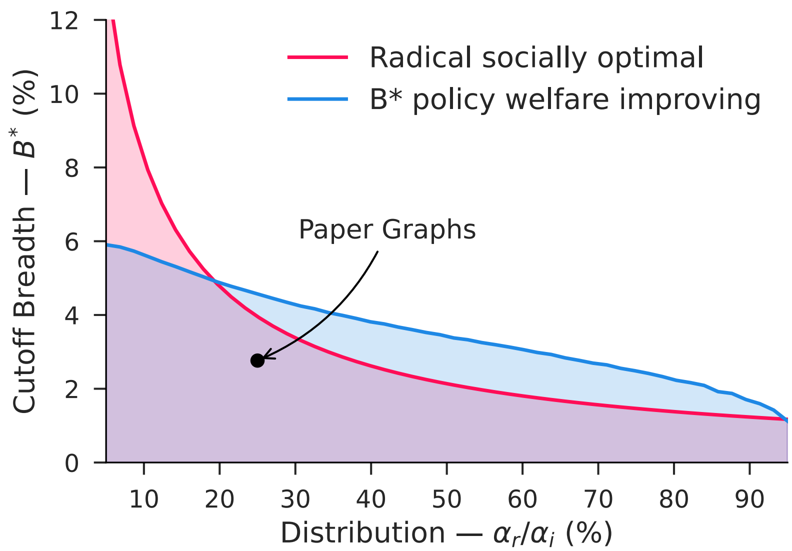


Figure 10: Sensitivy of breadth policy outcome to radical parameters (note that conditional on α_r, the cutoff breadth B^* is monotone in η_r)

novation 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.

Even using logarithms, the scaling of Figure 9 is somewhat constrained. As such, we find it useful to consider an alternative parameterization. In Figure 10, we keep the relative value of α_r on the x-axis but we plot the value of B^* on the y-axis. Note that this works because B^* is monotone in the cost parameter η_r. A higher value of B^* means that one must set a larger breadth in order to induce radical innovation, which corresponds to a higher cost for radical innovation, meaning a lower value of η_r. Thus by focusing on the case where $B^* \geq 1$, we implicitly only consider parameter values for which radical innovation does not arise in equilibrium and must be induced with patent 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 cases 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 80%, while the optimal subsidy once the switch to radical innovation is made falls to roughly 30%. 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 very often optimal to induce.

6 Policy Implications and Conclusions

6.1 Achieving optimal patent breadth

Incentives for substantive innovation are likely to remain weak while infringement damages are low. An order of magnitude increase in damage infringement fees for invention patents across the entire damage award distribution is probably necessary to achieve this needed level of infringement deterrence. However, it may be necessary to phase this increase in over time, for reasons discussed below. The distribution of damage awards for utility models and design patents should not substantively change from current levels. The substantial gap between damage awards that thereby emerges should tilt incentives for patent applicants against utility models or design patents (which undergo no real review for novelty) and in favor of invention patents, whose quality evaluation should be substantially strengthened in the manner described below.

However, even with these changes, some patent disputes will still need to be settled in court. Current incentives that reward judges for speedy resolutions that rely on statutory default damage levels should be replaced with incentives for careful jurisprudence that establishes legal precedents, especially those which reinforce the national goal of higher patenting standards with respect to novelty and inventive step. Courts should be incentivized to use expert witnesses appropriately, since the patent bar currently lacks the expertise to define patent boundaries in consistent ways, especially in complex and quickly evolving technical domains. Legal regulations should be changed to allow foreign legal and subject matter experts to act as expert witnesses and to provide counsel and advice to civil litigants, especially plaintiffs in infringement cases. These liberalizations should extend to an expanded role for international law firms in Chinese civil litigation, at least with regard to intellectual property disputes. In addition, Chinese courts should admit as evidence the results of foreign patent evaluations and judgments. All of these changes will speed up the process by which Chinese patent courts can converge to international best practices. Changes in legal procedure and standards will inevitably cause delays in case resolution and an increase in uncertainty of outcomes, at least in the short run. This should be viewed as the price of progress. All of these changes should increase the ability of the Chinese patent system to lower the currently high level of statistical noise in the patent evaluation process and more consistently enforce what amounts to a higher "B" - a higher standard for novelty/inventive step – in China's intellectual property system.

6.2 Promoting more rigorous patent examination and increasing patent fees

The ability of the Chinese patent office to evaluate patent quality needs to be increased. This goal can be accomplished by reducing the number of low-quality applications and by increasing the resources devoted to patent evaluation and the rigor of the evaluation process. The abolition of direct patent subsidies should help limit the flow of low-quality applications (see Appendix B for details on China's patent subsidies). Indirect patent subsidies should also be reduced by replacing the statutory criteria in the laws establishing programs like Innocom that currently emphasize patent quantity with criteria that emphasize patent quality and/or non-patent measures of research quality. A significant increase in invention patent application and renewal fees could usefully rebalance the incentives to file low quality patent application; the revenues generated by these (much) higher fees could increase the resources within the patent office to evaluate patents applications and could support post-grant reviews of low quality patents. Similarly, utility model application and renewal fees should be sharply increased to ameliorate the explosion of utility models, and applicants should be forced to choose at the time of initial application between a utility model or an invention patent. They should not be allowed to defer this choice until an invention patent issues. This alone is likely to reduce the incentive for firms to take out multiple utility models for every invention patent. Government directives that set formal or informal patent application and grant targets for regions, industries, or firms should be ended and replaced with government directives that explicitly target a relative decline in low quality or incremental patent applications and grants.

As already noted, some portion of the additional revenues generated by substantially higher application and renewal fees should be directed (by statute) to support the expansion and acceleration of China's post-grant patent review system. The current practice of Chinese patent litigation proceedings pausing while litigated patents undergo post-grant review should be encouraged through procedural pricing (i.e., keeping the price of review low relative to the price of civil litigation). New legislation should impose a relatively high burden of proof on the parties in this post-grant review procedure seeking to defend the validity of their patents. A well-resourced post-grant review system could play an important role in the winnowing of the patent thicket and in the reinforcement of higher standards for inventive novelty throughout the patent system.

6.3 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 countries 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. To reach its innovative potential, China needs fewer but better patents, (much) higher damage awards, higher (net) patent fees, and greater discrimination between invention patents and utility models. Will these necessary reforms be adopted? Only time will tell.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., & Kerr, W. (2018). Innovation, reallocation, and growth. *American Economic Review*, *108*(11), 3450–91.

Acemoglu, D., Akcigit, U., & Celik, M. A. (2020). Radical and incremental innovation: The roles of firms, managers and innovators. *AEJ Macroecon (forthcoming)*.

Agarwal, R., Ganco, M., & Ziedonis, R. H. (2009). Reputations for toughness in patent enforcement: Implications for knowledge spillovers via inventor mobility. *Strategic Management Journal*, *30*(13), 1349–1374.

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*, 323–351.

Akcigit, U., Hanley, D., & Serrano-Velarde, N. (2021). Back to basics: Basic research spillovers, innovation policy, and growth. *The Review of Economic Studies*, *88*(1), 1–43.

Akcigit, U., & Kerr, W. R. (2018). Growth through heterogeneous innovations. *Journal of Political Economy*, *126*(4), 1374–1443.

Arora, A., Fosfuri, A., & Gambardella, A. (2004). *Markets for technology: The economics of innovation and corporate strategy*. MIT press.

Arora, A., & Gambardella, A. (1994). The changing technology of technological change: General and abstract knowledge and the division of innovative labour. *Research policy*, *23*(5), 523–532.

Azoulay, P., Fons-Rosen, C., & Graff Zivin, J. S. (2019). Does science advance one funeral at a time? *American Economic Review*, *109*(8), 2889–2920.

Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar extinction. *The Quarterly Journal of Economics*, *125*(2), 549–589.

Barry, C., Arad, R., Ansell, L., Cartier, M., & Lee, H. (2017). 2017 patent litigation study: Change on the horizon?

Belderbos, R., Park, J., & Carree, M. (2021). Do r&d investments in weak ipr countries destroy market value? the role of internal linkages. *Strategic Management Journal*, *42*(8), 1401–1431.

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

Beukel, K., & Zhao, M. (2018). Ip litigation is local, but those who litigate are global. *Journal of International Business Policy*, *1*(1), 53–70.

Branstetter, L. G., Fisman, R., & Foley, C. F. (2006). Do stronger intellectual property rights increase international technology transfer? empirical evidence from us firm-level panel data. *The Quarterly Journal of Economics*, *121*(1), 321–349.

Branstetter, L. G., & Sakakibara, M. (2002). When do research consortia work well and why? evidence from japanese panel data. *American Economic Review*, *92*(1), 143–159.

Byun, S. K., Oh, J.-M., & Xia, H. (2021). Incremental vs. breakthrough innovation: The role of technology spillovers. *Management Science*, *67*(3), 1779–1802.

Chen, Y., Pan, S., & Zhang, T. (2018). Patentability, r&d direction, and cumulative innovation. *International Economic Review*, *59*(4), 1969–1993.

Chen, Z., Liu, Z., Suárez Serrato, J. C., & Xu, D. Y. (2021). Notching r&d investment with corporate income tax cuts in china. *American Economic Review*, *111*(7), 2065–2100.

Chinese inventiveness shows the weakness of the law. (2020). *The Economist*, *January 4th*.

Cornelli, F., & Schankerman, M. (1999). Patent renewals and r&d incentives. *The RAND Journal of Economics*, 197–213.

Dang, J., & Motohashi, K. (2015). Patent statistics: A good indicator for innovation in china? patent subsidy program impacts on patent quality. *China Economic Review*, *35*, 137–155.

Dvorak, P., & Woo, S. (2019). The world expected a chinese tech takeover. alibaba can't even conquer vietnam. *The Wall Street Journal*, *September 9th*.

Fitzgerald, T., Balsmeier, B., Fleming, L., & Manso, G. (2021). Innovation search strategy and predictable returns. *Management science*, *67*(2), 1109–1137.

Ganco, M., Ziedonis, R. H., & Agarwal, R. (2015). More stars stay, but the brightest ones still leave: Job hopping in the shadow of patent enforcement. *Strategic Management Journal*, *36*(5), 659–685.

Gilbert, R., & Shapiro, C. (1990). Optimal patent length and breadth. *The RAND Journal of Economics*, 106–112.

Grossman, G. M., & Helpman, E. (1991). Quality ladders in the theory of growth. *The review of economic studies*, *58*(1), 43–61.

Hall, B. H., & Ziedonis, R. H. (2001). The determinants of patenting in the us semiconductor industry, 1980-1994. *Rand Journal of Economics*, *32*(1), 101–128.

Helmers, C. et al. (2018). *The economic analysis of patent litigation data* (Vol. 48). WIPO.

Hopenhayn, H. A., & Mitchell, M. F. (2001). Innovation variety and patent breadth. *RAND Journal of Economics*, 152–166.

Hu, A. (2014). Propensity to patent, competition and china's foreign patenting surge. *Intellectual property for economic development*. Edward Elgar Publishing.

Hu, A. G., & Jefferson, G. H. (2009). A great wall of patents: What is behind china's recent patent explosion? *Journal of Development Economics*, *90*(1), 57–68.

Hu, W., Yoshioka-Kobayashi, T., & Watanabe, T. (2020). Determinants of patent infringement awards in the us, japan, and china: A comparative analysis. *World Patent Information*, *60*, 101947.

Jensen, R., & Thursby, M. (2001). Proofs and prototypes for sale: The licensing of university inventions. *American Economic Review*, *91*(1), 240–259.

Kennedy, S. (2017). *The fat tech dragon: Benchmarking china's innovation drive*. Center for Strategic & International Studies.

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

Lamin, A., & Ramos, M. A. (2016). R&d investment dynamics in agglomerations under weak appropriability regimes: Evidence from indian r&d labs. *Strategic management journal*, *37*(3), 604–621.

Lerner, J. (2009). The empirical impact of intellectual property rights on innovation: Puzzles and clues. *American Economic Review*, *99*(2), 343–48.

Li, X. (2012). Behind the recent surge of chinese patenting: An institutional view. *Research policy*, *41*(1), 236–249.

Lovely, M. E., Liang, Y. et al. (2018). Trump tariffs primarily hit multinational supply chains, harm us technology competitiveness.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science*, *2*(1), 71–87.

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

Nordhaus, W. D. (1969). An economic theory of technological change. *The American Economic Review*, *59*(2), 18–28.

O'Donoghue, T., Scotchmer, S., & Thisse, J.-F. (1998). Patent breadth, patent life, and the pace of technological progress. *Journal of Economics & Management Strategy*, *7*(1), 1–32.

Paik, Y., & Zhu, F. (2016). The impact of patent wars on firm strategy: Evidence from the global smartphone industry. *Organization Science*, *27*(6), 1397–1416.

Qian, Y. (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*(3), 436–453.

Sun, Z., Lei, Z., Wright, B. D., Cohen, M., & Liu, T. (2021). Government targets, end-of-year patenting rush and innovative performance in china.

Thursby, J. G., & Thursby, M. C. (2002). Who is selling the ivory tower? sources of growth in university licensing. *Management science*, *48*(1), 90–104.

Wei, S.-J., Xu, J., Yin, G., & Zhang, X. (2021). Bureaucratic bean counting and patent subsidies: A welfare evaluation of china's pro-innovation program.

Yamaguchi, K. (2010). Japanese patent litigation and its related statistics—current environment and future agenda—. *Notes*, *2011*(2012), 2013.

Zhao, M. (2006). Conducting r&d in countries with weak intellectual property rights protection. *Management science*, *52*(8), 1185–1199.

Appendices

A Appendix Tables

Table A1: Calibration of inventive steps for incremental versus radical innovation

	Min	1%	5%	10%	25%	50%	75%	90%	95%	99%	Max	Mean	s.d.	#obs
Family patents (registered in the U.S.) of Chinese patents made by domestic Chinese firms	1	1	1	1	2	5	11	22	34	88	528	10.05	18.13	18068
All patents (registered in the U.S.)	1	1	1	1	3	6	15	36	61	170	3724	16.81	48.63	1133267
All patents (registered in the U.S., self citations excluded)	1	1	1	1	2	5	14	32	56	159	3577	15.4	45.59	1094413
Chinese patents with family patents filed by domestic Chinese firms	1	1	1	1	2	4	7	13	18	30	247	5.68	6.69	19563
Chinese patents without family patents filed by domestic Chinese firms	1	1	1	1	2	3	6	11	14	25	370	4.8	5.42	158560

Table A2: Patent production function estimates on 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 RD expenditures as well as positive number of patents. Patents refer to invention patents only, counted by their filing years. 1-SD for RD expenditures, family patents and non-family patents are 208291.9 (1k RMB), 50, 86.6, respectively. Robust standard errors are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3: Alternative estimation of R&D cost of radical and incremental innovations

Dependant variable	**R&D spending**			
	(1)	(2)	(3)	(4)
Samples	R&D manufacturing firms	R&D manufacturing firms	publicly traded firms	publicly traded firms
#newinIPC3	0.0713	0.0383	0.0238**	0.0307**
	(0.0527)	(0.0438)	(0.0107)	(0.0134)
#nonNewinIPC3	0.734***	0.715***	0.337***	0.237***
	(0.0688)	(0.0738)	(0.0230)	(0.0276)
cost of radical innovation	**18334.4**		**84126316**	
cost of incremental innovation	**1163.491**		**3950454**	
#employees		0.133		1.183***
		(0.142)		(0.231)
assets/debts		0.00894		-0.00141
		(0.0275)		(0.0199)
ROA		1.771		0.0350
		(1.988)		(0.0276)
intangible/assets				-0.0212
				(0.0402)
2-digit industry FEs	Y	Y	Y	Y
Constant	0.0220	0.548	-0.0994***	0.0764
	(0.0546)	(2.174)	(0.0194)	(0.0756)
Observations	2,404	2,403	1,096	496
R-squared	0.662	0.686	0.441	0.826

Note:Standardized estimates report the beta coefficients. Robust standard errors are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A4: Distribution of forward citations received by invention patents made by domestic Chinese firms by year

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).

B Proofs

B.1 Baseline Model

Proof that $\Lambda/\Omega \leq 1$. Applying Jensen's inequality for the log function, we find

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

Exponentiating this yields the desired result. □

B.2 Imperfect Examination

The equation describing the expected return of a new innovator, $\tilde{z}_k$ will now be

$$\tilde{z}_k = \frac{\left(\frac{1}{1+\alpha_k}\right) P_k}{\rho + \tau} \quad \text{where} \quad P_k = \varepsilon q + (1-\varepsilon)B^{-\alpha_k}$$

Following the same staps as in the baseline, the the choice of innovation type will depend on the relative values

$$H_i = \frac{\eta_i P_i}{1+\alpha_i} \quad \lesseqgtr \quad \frac{\eta_r P_r}{1+\alpha_r} = H_r$$

And the analogous relative distortion factor is

$$G(\alpha_k, B) \equiv \left(\frac{\alpha_k}{1+\alpha_k}\right)\left[\varepsilon q + (1-\varepsilon)B^{-\alpha_k}\right]$$

With $\varepsilon = 0$, we are in the baseline case. As ε grows larger, we will see B^* increase, and it will tend towards $B^* \to \infty$ as $\varepsilon \to 1$.

Though the generic equation characterizing the equilibrium remains unchanged, we will have different expressions for P_k (given above) and for the markup aggregates Λ and Ω, namely

$$\Lambda^{-1} = \left(\frac{\alpha_k}{1+\alpha_k}\right)\left[\frac{\varepsilon q + (1-\varepsilon)B^{-\alpha_k} \cdot B^{-1}}{\varepsilon q + (1-\varepsilon)B^{-\alpha_k}}\right]$$

$$\log(\Omega) = \frac{1}{\alpha_k} + \frac{(1-\varepsilon)B^{-\alpha_k}}{\varepsilon q + (1-\varepsilon)B^{-\alpha_k}} \cdot \log(B)$$

B.3 Patent Fees

The resulting type-conditional expected gain from innovation is now

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

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

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

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 resulting innovation rates for the case where $B = B^*$, these pure regimes will arise when

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

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.

Let H^* denote the common value of H_k at the critical patent breadth B^* or

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

The free entry condition for both types of innovation evaluates to

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

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

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

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

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

Plugging this into the free entry condition, we ultimately find

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

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

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

A sufficient condition for this to hold 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 have

$$(\beta - s)(1+\alpha_r) > 1$$

Given a value for $\bar{\alpha}$, one can derive the implied weights f_i and f_r. Finally, we can compute the aggregate growth rate

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

where $\hat{\alpha}$ is the geometric mean of α, defined by

$$
\hat{\alpha}^{-1} = \frac{f_i}{\alpha_i} + \frac{f_r}{\alpha_r}
$$

B.4 Random Innovation Costs

Let the joint distribution of η_i and η_r be independently distributed Fréchet random variables with shape parameter σ and location parameters $\bar{\eta}_k$. In the general case, a Fréchet random variable with mean μ and shape parameter σ has the cumulative density

$$
H(x|\mu,\sigma) = \exp\left(-\left(\frac{x}{\mu}\right)^{-\sigma}\right)
$$

This results in a density function of the form

$$
h(x|\mu,\sigma) = \frac{\sigma}{\mu}\left(\frac{x}{\mu}\right)^{-\sigma-1} \exp\left(-\left(\frac{x}{\mu}\right)^{-\sigma}\right)
$$

Now suppose we have n i.i.d. Fréchet random variables with means μ_i and common shape parameter σ. Now consider, for a given i, the probability that i takes on value $x_i = x$ and all other $j \neq i$ satisfy $x_j \leq x$. Integrating this probability over $h(x|\mu_i,\sigma)$ yields the probability that i is the maximal element F_i. It is useful to define variable

$\bar{\mu} = [\sum_i \mu_i^\sigma]^{1/\sigma}$ and $z = \left(\frac{x}{\bar{\mu}}\right)^{-\sigma}$ then proceed as follows

$$\begin{aligned}
F_i &= \int_0^\infty h(x|\mu_i,\sigma) \prod_{j\neq i} H(x|\mu_j,\sigma)dx \\
&= \int_0^\infty \frac{\sigma}{\mu_i}\left(\frac{x}{\mu_i}\right)^{-\sigma-1} \exp\left(-\left(\frac{x}{\bar{\mu}}\right)^{-\sigma}\right) dx \\
&= \left(\frac{\mu_i}{\bar{\mu}}\right)^\sigma \int_0^\infty \exp(-z)dz \\
&= \left(\frac{\mu_i}{\bar{\mu}}\right)^\sigma = \frac{\mu_i^\sigma}{\sum_j \mu_j^\sigma}
\end{aligned}$$

In our case, we have $n = 2$ and $\mu_k = \bar{\eta}_k z_k$, so the proportion of firms choosing each type of innovation will be

$$F_k = \frac{(\bar{\eta}_k z_k)^\sigma}{(\bar{\eta}_i z_i)^\sigma + (\bar{\eta}_r z_r)^\sigma} = \frac{1}{1 + \left(\frac{\bar{\eta}_{-k}}{\bar{\eta}_k} \cdot \frac{z_{-k}}{z_k}\right)^\sigma}$$

And the efficiency result will carry over to this settings (since it is true for all η_k combinations), meaning if B is set appropriately, the fraction of each type of innovation will be the same in equilibrium as that of the social planner.

We will also want to know the realized step size distribution. To find this, we must calculate

$$\begin{aligned}
E_k &= \int_0^\infty xh(x|\mu_i,\sigma) \prod_{j\neq i} H(x|\mu_j,\sigma)dx \\
&= \int_0^\infty x\frac{\sigma}{\mu_i}\left(\frac{x}{\mu_i}\right)^{-\sigma-1} \exp\left(-\left(\frac{x}{\bar{\mu}}\right)^{-\sigma}\right) dx \\
&= \left(\frac{\mu_i}{\bar{\mu}}\right)^\sigma \bar{\mu} \int_0^\infty z^{-1/\sigma} \exp(-z)dz \\
&= \bar{\mu}\Gamma\left(\frac{\sigma-1}{\sigma}\right)\left(\frac{\mu_i}{\bar{\mu}}\right)^\sigma = \bar{\mu}\Gamma\left(\frac{\sigma-1}{\sigma}\right) F_k
\end{aligned}$$

And the overall growth rate will be given by

$$g = \Gamma\left(\frac{\sigma-1}{\sigma}\right) R \sum_k \frac{\bar{\mu} F_k}{z_k} \cdot P_k \ \mathbb{E}_k[\log(\gamma)|\gamma > B]$$

where one can show

$$P_k \ \mathbb{E}_k[\log(\gamma)|\gamma > B] = \int_B^\infty \log(\gamma) \cdot \alpha_k \gamma^{-\alpha_k - 1} \ d\gamma = B^{-\alpha_k}\left[\log(B) + \frac{1}{\alpha_k}\right]$$

So that we can define $\hat{\sigma} = \frac{\sigma-1}{\sigma}$ and use the fact that $F_k = (\bar{\eta}_k z_k / \bar{\mu})^\sigma$ to get

$$g = \Gamma(\hat{\sigma}) \left(F_i^{\hat{\sigma}} g_i + F_r^{\hat{\sigma}} g_r \right)$$

where

$$g_k = \bar{\eta}_k R B^{-\alpha_k} \left[\log(B) + \frac{1}{\alpha_k} \right]$$

Notice that when $\sigma = 1$, we get $g = g_i + g_r$.

In the case of the social planner, they will solve this optimization at each time period and the outcome will be equivalent to the case where $B = 1$ and $z_k = 1/\alpha_k$.

We also need to think about the ex ante free entry condition. Here, we have

$$E = E_r + E_r = \Gamma(\hat{\sigma})\bar{\mu}$$

and the free entry condition is $E = w$ implying $\tilde{E} = \tilde{w}$. The determination of $\tilde{w}$ depends on the λ distribution. This will be a mixture of the r and i distribution with weights

$$W_k = \frac{F_k P_k}{F_r P_r + F_i P_i}$$

The contibution from each component will be as before: $\Lambda_k = \frac{\alpha_k}{\alpha_k + 1}$, so that

$$\Lambda^{-1} = W_r \Lambda_r + W_i \Lambda_i = \frac{F_r P_r \Lambda_r + F_i P_i \Lambda_i}{F_r P_r + F_i P_i}$$

With repeated substitution, this yields a closed form expression for R, although, one can see that Λ^{-1} will not depend on R due to the common $\rho + \tau$ term cancelling. This term will also factor linearly out of $\bar{\mu}$. Using a similar derivation to the growth rate, we find

$$\tau = \Gamma(\hat{\sigma})(F_r^{\hat{\sigma}} \tau_r + F_i^{\hat{\sigma}} \tau_i) \quad \text{where} \quad \tau_k = \bar{\eta}_k P_k R$$

Thus the free entry condition will be of the form

$$\frac{\overline{\eta\pi}}{\rho + \bar{\eta} R} = \frac{\Lambda^{-1}}{1 - R}$$

and can hence be solved in closed form.

B.5 Risk Free Incremental

One complicating factor is that radical and incremental innovation step sizes in this model have both different means and different variances. In their paper, Y. Chen et al. (2018) instead focus on a dichotomy between safe and risky innovation. We can render our model somewhat comparable to that setting by replacing the incremental innovation technology with a risk-free outcome that always yields step size λ_i, while the radical side is still Pareto distributed with parameter α_k.

We ensure that the two innovation types differ only in their risk profile be equating the expected profit arising from each in a zero breadth ($B = 1$) setting

$$\frac{1}{1+\alpha_r} = 1 - \lambda_i^{-1}$$

while continuing to allow different cost levels η_i and η_r. In this case, the patent breadth that induces radical innovation will satisfy

$$\frac{\eta_r B^{-\alpha_r}}{1+\alpha_r} = \eta_i(1-\lambda_i^{-1}) = \frac{\eta_i}{1+\alpha_r}$$

Resulting in a threshold value

$$B^* = \left(\frac{\eta_r}{\eta_i}\right)^{\frac{1}{\alpha_r}}$$

Note that we can trivially induce radical innovation by setting $B > \lambda_i$, so in order for this to operative, we must also assume

$$B^* < \lambda_i \quad \Rightarrow \quad \eta_r < \eta_i \lambda_i^{\alpha_r}$$

Thus we can see that the same approximate dynamic still holds when we restrict the differences between incremental and radical outcomes to be only risk related.

B.6 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. It is useful to derive the following variance decomposition first

$$\begin{aligned}
\mathbb{V}[c] &= \mathbb{E}[c^2] - \mathbb{E}[c]^2 \\
&= \mathbb{E}[\mathbb{E}[c^2|\gamma]] - \mathbb{E}[\mathbb{E}[c|\gamma]]^2 \\
&= \mathbb{E}\left[\mathbb{V}[c|\gamma] + \mathbb{E}[c|\gamma]^2\right] - \mathbb{E}[\mathbb{E}[c|\gamma]]^2 \\
&= \mathbb{E}\left[\mathbb{V}[c|\gamma]\right] + \mathbb{E}[\mathbb{E}[c|\gamma]^2] - \mathbb{E}[\mathbb{E}[c|\gamma]]^2 \\
&= \mathbb{E}\left[\mathbb{V}[c|\gamma]\right] + \mathbb{V}[\mathbb{E}[c|\gamma]]
\end{aligned}$$

We will also need the following properties of the Pareto distribution

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

Our particular assumption for the citation data generating process is that both conditional mean and variance are linear in the step size, which would be consistent with a Poisson distribution

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

The unconditional mean of citations is then straightforward to calculate

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

For the unconditional variance, we use the equation derived above

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

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 suffer from the issue of depending on κ, which is a nuisance parameter in our case. For instance, if one calculates the typical index of disersion, we find

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

Alternatively, if one calculates the coefficient of variation, we arrive at th equally problematic

$$C = \frac{\sqrt{\mathbb{V}[c]}}{\mathbb{E}[c]} = \sqrt{\frac{\alpha - 1}{\kappa} + \frac{\alpha}{\alpha - 2}}$$

Thus we use the modified index of dispersion ($\tilde{D}$) for the calibration.

C China's patent subsidies

We have shown that even though China's patent enforcement is not sufficient to protect valuable inventions, China has nevertheless witnessed an intensive patent explosion in recent years. Why are Chinese patent applicants still enthusiastically filing patents on their inventions when the weakness of patent enforcement does not protect them from potential infringement? Though prior studies (A. G. Hu and Jefferson, 2009; A. Hu, 2014) have identified several factors driving China's patent explosion such as economic growth, increasing exports, and FDI, strengthened institutions, we argue that the prevalence of patent subsidies is the most important factor that explains the paradox of a patent explosion under a weak-appropriability regime.

Two types of patent subsidies are of our interest. One is the direct patent subsidy, designed to induce higher levels of patenting by offsetting a variety of patenting fees. These direct subsidies are mostly provided by local governments and date back as early as 1999, when Shanghai began compensating patent applicants in order to boost patent filings. After that, patent subsidy programs were quickly adopted by many other local governments. By 2007, 97% of local governments at the province level had followed suit (Li, 2012). Though procedures and the amount of subsidies vary across regions, the majority of direct patent subsidies reimburse patent filing fees, whether the patent application is ultimately approved or not. Half of subsidy programs reimburse examination expenses for invention patents and provide awards for relatively higher-quality patents such as invention patents grants and PCT applications. The subsidy program offered by the city Guangzhou, capital of Guangdong province, provides an example of the array of incentives generated by these programs. The city of Guangzhou began its subsidy program in 2002, providing firm applicants 3500 RMB ($ 443) for each individual invention application, 5000 RMB ($604) for each PCT application, and 10,000 RMB ($1208) for direct foreign patent filing, but only 400 RMB ($ 48) per granted utility model. Subsidies for non-firm applicants are similar but slightly lower. By comparison, in 2001 the application fee for an invention patent was 1850 RMB ($234); the application fee for utility model was 500 RMB ($60); and the application fee for a PCT application before entering into the national phase of examination was around 15000 RMB ($1900). If application fee discount is applied, the application fee will be lower by 15% to 30%. Therefore, the bulk of patent application fee is covered by the subsidy; in the case of invention applications or application fee discount, filing a patent could even earn money for the applicant.

Several studies have shown that direct patent subsidy programs boost patent applications in China. As one of the first efforts in this line of research, Li (2012) shows that the province-level patent subsidy programs induced an increase in patent propensity among firms, universities, research organizations and individuals. Dang and Motohashi (2015) take a further step in investigating the effect of direct patent subsidies at the firm level. By examining the effect of provincial subsidies on middle- and large-size manufacturing companies, Dang and Motohashi further corroborate Li (2012) finding and show that apart from R&D input and financial outputs, patent subsidies are another key driver of the patent explosion, contributing to around 30% of the total increase in patent filing. They also cast doubt on the quality of these subsidy-driven patent applications, raising concerns for those who intend to use patent counts as a measure of innovative output. Partly in reaction to these studies, the Chinese central government has sought to wind down local government's direct subsidy programs since 2020. China's patent office issued a directive to local governments to end all patent subsidies for patent applications by June of 2021[10].

If actually implemented, cancellation of direct patent subsidies on patent applications might partly slow down China's patent explosion and mitigate some problems caused by that explosion, including the growing challenge of defining patent boundaries in China's ever denser patent thicket. But policy-induced patent filing will not entirely evaporate, even if direct subsidies disappear, because indirect patent subsidies, the other type of patent subsidy, is still playing a key role in China's innovation landscape. We define indirect patent subsidies as government subsidies which are contingent on application or granted ownership of patent rights but are not directly funding patent applications or grants. A representative example of this type of subsidy, which is probably the first indirect patent subsidy program at the central government level, is China's InnoCom policy. Issued in the early 1990s, the InnoCom policy provides high-tech certificates for firms which meet a set of qualifying criteria based on firms' innovation capacity in high-tech areas. The InnoCom certificate brings about a variety of benefits to qualified high-tech firms, including but not limited to a reduction of the firm's corporate income tax rate from 33% (or 25% after 2008) to 15%, direct cash awards from local governments, favorable provisions for land usage, etc. The criteria for certificate qualification were changed in 2008, adding patent acquisition as a key performance metric: ownership of patents became a necessary condition for certification patent right ownership is a veto to certification; and the more patents rights the firm owns, the more likely it will win the InnoCom certification, if other

[10]Despite the central government's orders to eliminate subsidies for patent applications, most local governments have merely made small adjustments to their subsidy programs. As a consequence, subsidies for patent grants and annual renewal fees for granted patents are still in force in most places. See http://banshi.beijing.gov.cn/pubtask/task/1/110000000000/8e01aa7c-1795-4fe8-835e-9813f753d468.html?locationCode=110000000000 for Beijing's latest patent subsidy policy

criteria are fulfilled. The InnoCom policy might incentivize firms to patent truly valuable technologies which might have not been patented without InnoCom. But firms are also incentivized by this new "patent" rule imposed by InnoCom, probably to a larger extent, to patent low-value inventions such as utility models, just to fulfill the InnoCom criteria. The latter patenting incentive may further fuel China's low-quality patent explosion.

Inspired by the InnoCom policy, an increasing number of new indirect subsidy policies, both from the central government and the local government, include patent applications (and sometimes patent grants) as an important evaluation metric for firms' innovation capacity. Exampls include the Torch high-tech industry development program and the InnoFund program. Though systematic studies on the effect and implication of these indirect subsidy policies are still rare, it is reasonable to hypothesize that firms, other organizations, or even individuals who rely on patents to be qualified for certain certificates or benefits will be incentivized to file patents which would not have been filed without this IP requirement. Similar to the "extra" patent filings driven by direct patent subsidy programs, indirect patent subsidy programs tend to generate patented inventions with limited value.

D Alternative approach of estimating innovation costs

In this part, we lay out details of estimating the costs of making incremental innovation vis-a-vis radical innovations.

To start with, we compile a data base comprising of Chinese manufacturing companies, which tend to have fewer diversified product lines and smaller size. We begin with Chinese Annual Survey of Manufacturing (ASM), a longitude survey conducted by Bureau of Statistics which is widely used by scholars. We then select companies with positive R&D records from 2005 to 2007, a sub-sample of firms that are of particular interest to our inquiry. Then We match these firms with their patent records from IncoPat. The matching is based on firms' full names and addresses; for patent applicants whose unique IDs [11] are available in the Incopat database, we also match on these unique identification numbers. We employ a multivariate regression to control other possible contributors to R&D spending. Specifically, we estimate the following specification:

[11] A unique firm identification number issued by the government for tax purpose.

$$Y_i = \alpha + \beta_1 R\&D_i + \beta_2 X_i + \delta_j + \epsilon_i$$

Where Y_i is the number of family patents (or non-family patents) filed by firm i from 2005 to 2007 [12]. $R\&D_i$ are the R&D expenditure of firm i during the same time period. X_i are a battery of control variables, including the number of employees, ROA (returns to assets) and the ratio of total assets over total debt. We also insert industry fixed effects, δ_j, to control unobserved time-invariant factors of the industry j in which firm i belongs to [13]. All variables are aggregated over the entire observation window, years from 2005 to 2007. Through this specification, we could estimate the number of family patents (or non-family patents) that are produced when the firm adds one additional unit of R&D expenditures.

Table A2 reports the results. As a robustness check, we try out different forms of variables: standardized variables in columns 1 and 2; logarithm transformed variables in columns 3 and 4; raw numbers in columns 5 and 6. The signs and statistical significance of coefficients of R&D spending are similar across different forms of variables. Particularly, we focus on results from specifications using standardized variables and raw numbers since the log-log specification provides elasticity estimates (thus more suitable for estimating the entire distribution). Columns 1 and 2 indicate that on average the per-patent R&D cost is around 5503090 RMB and 2594630 RMB for family patents and non-family patents, respectively (row 2, figures in bold). Estimates from specifications using raw numbers are similar. As expected, the estimates on R&D costs from the multivariate regressions turn out to be much lower than those estimates from the "naive" approach. These results indicate that the ratio of family patents- proxy for radical innovations- over non-family patents- proxy for incremental innovations- is around 2.12.

One might concern that the data we use for this approach might only reflect situation more than 10 years ago. Therefore, we try out 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 consequent years. Specifically, we collect all publicly traded companies with positive R&D inputs from 2014 to 2019 from CSMAR, a dedicated data set comprising information on publicly traded firms in China [14]. Following

[12] We restrict our observation window in this way because the R&D spending data is only available during this time period

[13] If the firm belongs to more than one industries during the observation period,

[14] We focus on this time period to get abundant data points. The R&D spending data of publicly traded firms is only available after 2007. Around 80% publicly traded firms lack R&D spending data

this, we match these publicly traded companies with their patenting records using their unique trading number in the stock market; the patenting data is from InnoPat. Finally, we exclude firms with zero patent filings from 2014 to 2019 because patents are proxies of innovation output in our calibration. This yields a data set comprising of 1098 publicly traded firms with 88 of them registered at least one Chinese patent oversea (firms making radical innovations according to our previous definition) and 1010 firms without any domestic patents being sought patent protection oversea (i.e. incremental innovators). Aside from regressing the patent output on R&D spending like what we have done in Table A2, here we regress in the opposite direction. Specifically, we regress the R&D spending on the number of radical innovations and the number of incremental innovations. In addition, we employ another definition of radical versus incremental innovations, which have been widely utilized in the literature (Byun et al., 2021; Fitzgerald et al., 2021). Specifically, we count the number of patents filed in technological area (3-digit IPC) where the focal firm has not filed patents yet; the measure is counted in each year as a proxy for the radical innovations whereas patents filed in tech areas where the focal firm already has patents before are defined as incremental innovations. We regress companies' R&D spending on the number of incremental and radical innovations they make and infer the the ratio of innovation costs of these two types of innovations from the regression coefficients. Formally, we estimate:

$$R\&D_i = \alpha + \beta_1 Radical_i + \beta_2 Incremental_i + \beta_3 X_i + \delta_j + \epsilon_i$$

Where $R\&D_i$ denotes R&D spending of firm i. $Radical_i$ and $Incremental_i$ are the number of radical innovations and incremental innovations filed by firm i. X_i are a battery of control variables at the firm level, including the number of employees, ROA (returns to assets), the ratio of total assets over total debt and the ratio of intangible assets over total assets (only applicable for publicly traded firms). δ_j is industry fixed effects.

for years from 2007 to 2013.