Haste Makes Waste? Quantity-Based Subsidies under Heterogeneous Innovations *

Linyi Cao Helu Jiang Guangwei Li Lijun Zhu †

September 5, 2023

Abstract

With quantity-based innovation targets and subsidy programs launched since the mid-2000s, China has seen a patent surge, accounting for 46% of the world's total patent applications in 2020; however, the overall patent quality has been declining after 2008. This paper develops a Schumpeterian growth model featuring innovating firms' quantity-quality trade-off between radical and incremental innovations, and decomposes subsidies' aggregate impact into quantity and quality channels. We calibrate the model to Chinese firm-level data in the early 2010s. Our quantitative analysis shows that the quality channel effects are negative and dominant, and quantity-based subsidies in that period reduce the TFP growth rate and welfare by 0.19 percentage points and 3.31%, respectively. We evaluate welfare gains under a constrained planner's problem, and propose skill subsidies which are quality-biased and effectively recover the optimal allocation.

JEL Classification: O31, O38, O40

Keywords: Heterogeneous Innovations, Quantity-Based Subsidies, Human Capital

*We gratefully acknowledge helpful comments from Ufuk Akcigit, Michele Boldrin, Xiaomeng Ren, Yongseok Shin, Zheng Michael Song, Ping Wang, Shang-Jin Wei, Jianhuan Xu, Zhiwei Xu, Ge Yin, Bo Zhao, and seminar and conference participants at the SUFE Macro Workshop, EFG Meeting 2021, AMES 2022 & 2023, CICM 2023, ABFER 10th Annual Conference, ISFTE 2023, Annual SAET Conference 2023, and XiangjiangRiver Forum 2023. All remaining errors are ours.

†Cao: Institute for Advanced Research, Shanghai University of Finance and Economics. Email: caolinyi@sufe.edu.cn; Jiang: Institute for Advanced Research, Shanghai University of Finance and Economics. Email: jianghelu@sufe.edu.cn; Li: School of Entrepreneurship and Management, ShanghaiTech University. Email: ligw@shanghaitech.edu.cn; Zhu: Institute of New Structural Economics, Peking University. Email: lijunzhu@nsd.pku.edu.cn.

1 Introduction

Subsidies are widely used to stimulate innovation. This paper studies the impact of innovation subsidies in China since the middle 2000s, when, partly due to the fear of falling into the "middle-income trap", the Chinese government launched a series of initiatives to ensure the country's success in transiting to an innovation-oriented economy (Ding and Li, 2015). We first document that innovation targets set by the central and local governments are largely quantity-based. The quantity of patents, in particular, has been extensively adopted as a concrete indicator of innovation achievement. Under a large scale of innovation subsidies to help achieve these targets, China's invention patent applications have increased from slightly above 10 thousand in 1990, or 1.08% of the world's total, to around 1.5 million in 2020, accounting for 45.69% of the global total, raising concerns about the underlying patent quality.

To study the macro impact of such subsidies from micro-level incentives, we collect a panel of innovating firms from 1998 to 2013 and use the information on forward citations to classify patents into high-quality radical ones and low-quality incremental ones. Consistent with anecdotal evidence, we find a robust decline in the share of radical patents after the mid-2000s, in terms of the absolute level, and especially compared to a clear rising trend before the mid-2000s. In addition, the relative quality of incremental patents to radical ones, measured as the ratio of the average number of forward citations received by the former to the latter, displays a clear decline after the mid-2000s, suggesting a crowding effect that as more incremental innovations are pursued, their average impact decreases.

Then we develop a structural growth model featuring innovating firms' endogenous choices between radical and incremental innovations to quantitatively study the impact of quantity-based subsidies. Our model builds on Schumpeterian models with heterogeneous innovations (Akcigit and Kerr, 2018). Radical innovations significantly impact productivity, while incremental innovations build on existing radical ones and make marginal improvements, with their impact gradually diminishing as more incremental innovations are pursued. Different from existing works in which radical and incremental innovations are random outcomes, we introduce a scarce R&D resource and allow the pursuit of different kinds of innovations to be endogenous, which generates a micro-level quantity-quality trade-off between radical and incremental innovations. Motivated by empirical patterns, we further assume that radical innovations are more skill-intensive in that a larger proportion of skilled labor is needed to realize one such invention.

Policymakers cannot precisely identify the quality of innovations and base their policies on the number of innovation outcomes, e.g., the number of patents. As firms face a trade-off between radical and incremental inventions, quantity-based subsidies encourage overall innovations but also bring an undesired shift of R&D efforts toward cheaper but incremental trials. Under a general equilibrium context, when all innovating firms expand their R&D expenditures, the skill premium also increases, further tilting firms' R&D efforts away from the more skill-intensive radical innovations. To dissect their growth and welfare implications, we decompose the impact of quantity-based subsidies into three channels, each corresponds to an empirical finding just mentioned: a positive *quantity* channel that the subsidies promote innovations and creative destruction; a negative *quality-composition* channel that quantity-based subsidies lower the aggregate weight on radical innovations; and a negative *quality-crowding* channel that more incremental trials reduce their average production value.

We then calibrate the theoretical model to moments of Chinese innovating industrial firms from 2011 to 2013. In particular, we use moments regarding radical and incremental patents in the data to discipline parameters related to radical and incremental innovations in the model, showing that the introduction of quantity-based innovation subsidies accounts for 29% of the quantity surge, 56% of the decline in the radical patent share, and 75% of the decline in the relative quality of incremental patents observed between the pre- and post-2008 periods. Although the quantity channel tends to enhance overall growth, the quality channels are much more dominant, especially the *quality-crowding* channel. Overall, quantity-based subsidies reduce the equilibrium growth rate by 0.19 percentage points, or 10% of the actual TFP growth decline from 2001-2007 to 2008-2014, and reduce the aggregate welfare by 3.31%.

China is still relatively scarce in innovative, skilled labor despite its fast economic catch-up. In 2018, 27% of the Chinese population between age 25 and 34 have completed tertiary education, which is much lower than in other major patent-holding economies. Within the model's framework, we further evaluate the impact of two alternative subsidies: education subsidy and skilled labor subsidy, which effectively recover the social planner's allocation. In the model, skill is acquired through formal education before a worker enters the labor market, and these subsidies raise the skill supply. Since radical innovations are more skill-intensive, increasing the supply of skilled labor substantially reduces the R&D cost of pursuing such inventions. Thus, in contrast to quantity-based innovation subsidies, the alternatives we propose are quality-biased — they significantly promote aggregate growth and welfare by improving both innovation quantity and quality.

Our paper highlights the importance of considering firms' endogenous responses in designing effective innovation policies. In that regard, the paper is related to three strands of literature. The first is the creative destruction literature with heterogeneous firms (Klette and Kortum, 2004; Akcigit and Kerr, 2018; Acemoglu et al., 2022). In a model of creative destruction, Ates and Saffie (2021) characterize a quantity-quality trade-off induced by financial frictions among entrants, while such trade-off in our model is on the intensive innovation margin. Our model builds on Akcigit and Kerr (2018), which develops a model in which firms pursue radical and incremental innovations randomly. As mentioned, the key deviation in our model is to endogenize this decision to capture the quantity-quality trade-off faced by innovating firms. We also incorporate heterogeneity of R&D input structure and human capital, a crucial dimension for innovation in developing countries.

Our work also relates to studies investigating China's R&D policies and patent surges (Hu and Jefferson, 2009; Fang et al., 2017; Ang et al., 2014). Li (2012) finds that local innovation subsidy programs help stimulate patent applications. Chen et al. (2019) find that subsidies positively impact incremental innovations but not radical ones. A few recent papers (Chen et al., 2021; Branstetter et al., 2023; Wei et al., 2023) take a more structural approach to examine China's innovation policies. Wei et al. (2023) studies the InnoCom program in a three-stage static framework and finds that it hurts welfare due to bureaucratic bean counting, and patent trade exacerbates that loss. Our model of creative destruction emphasizes firms' quantity-quality trade-off and the role of quality crowding. Branstetter et al. (2023) investigate China's patenting system, and they argue that narrow patent protection in China skews R&D efforts toward incremental innovations. Our paper addresses a similar issue but caused by innovation subsidies. König et al. (2022) study the impact of R&D misallocation in China. They find that a large subsidy might even reduce the growth rate as it distorts firms' imitation-innovation decisions. Subsidies may hurt growth in our framework but through a different channel.

Lastly, this paper is related to research on the role of human capital in innovation and economic growth, dating back to Nelson and Phelps (1966). In Vandenbussche et al. (2006), as innovation is more intensive in skilled labor than imitation, skilled labor significantly impacts growth when a country approaches the technology frontier. Akcigit et al. (2020) incorporate higher education policy into an endogenous growth model. They find that the impact of R&D subsidies can be strengthened if combined with higher education policies that alleviate financial constraints for the young. Our paper follows this line of research in emphasizing the input dimension of R&D and the importance of human capital and education in promoting innovation.

The rest of the paper is organized as follows. Section 2 provides institutional background and describes motivational facts, and Section 3 introduces the model. The quantitative analysis is the focus of Section 4. Concluding remarks are presented in Section 5.

2 Institutional Background and Motivational Facts

This section first provides an overview of quantity-based innovation targets set by China's central and local governments since the mid-2000s and the associated patent surge. We then construct firm-level panel data to study the decline in patent quality in recent years.

2.1 Institutional Background and Patent Quantity

China started to emphasize the importance of building an "innovation-oriented" economy in the mid-2000s. In 2006, the Chinese central government released the *Outlines of Medium and Long-term National Plan for Science and Technology Development (2006-2020)*, which pronounced the building of an innovative economy as a new national strategy (Ding and Li, 2015; König et al., 2022). One critical and specific metric in the documentation is that by 2020, the total number of granted invention patents by Chinese nationals rank top 5 globally.[1] In 2010, China National Intellectual Property Administration issued the *National Patent Development Strategy 2011-2020*, which explicitly set the following *quantity* targets:

[1]Other specific targets listed in the documentation include the following. By 2020, the share of total R&D expenditures in GDP will achieve 2.5% or more; the contribution of technological progress to economic growth will account for more than 60%; the dependence on foreign technology will reduce to less than 30%; the total number of forward citations of international scientific papers by Chinese nationals will rank top 5 globally.

1. The total number of invention patents will rank top 2 in the world, and total patents reach 2 million in 2015;
2. Invention patents per million population will increase by 100% in 2015 and by 300% in 2020;
3. At least 8% of above-scale industrial enterprises will apply for patents in 2015 and 10% in 2020.

With these documents released by the central government, many local governments have also made explicit targets on the number of patents. In Table A.1 in Appendix, we list several patent quantity targets set in the 2000s and 2010s in developed areas like Beijing and Shanghai, as well as in relatively less developed northeastern Heilongjiang province and southwestern Guizhou province.

To help achieve these targets, the central and local governments issued supportive policies to promote firms' innovation activities. To encourage patent filing, the State Intellectual Property Office issued the *Measures of Patent Fee Deferral* in 2006. Many local governments have since issued additional incentives for patenting (Ding and Li, 2015). For example, the Beijing city government subsidizes up to 2,150 Chinese Yuan (CNY) for an invention patent application. The Zhejiang provincial government grants each invention patent a one-time 3,000 CNY subsidy. By 2008, 29 of 32 provincial governments have introduced patent subsidy programs in mainland China (Li, 2012).[2]

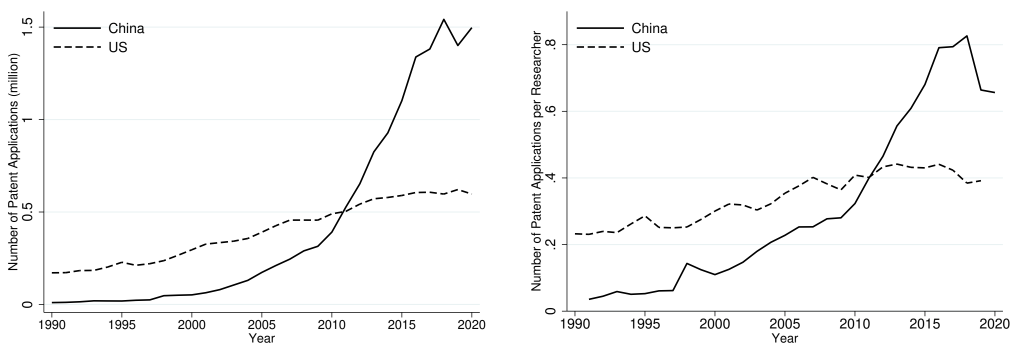
With the explicit quantity targets and associated subsidy policies, China has seen a dramatic surge in the total number of invention patents.[3] Figure 2.1 presents the evolution of the total number of newly applied patents (panel (a)) and patents per researcher (panel (b)). China's total number of patent applications in the 1980s and 90s was substantially smaller than the US. It accounted for 1.02% of the world's patent applications in 1990. In

[2] Another widely adopted policy is the intellectual property rights pledge financing (Ding and Li, 2015).

[3] There are three types of patents in China: invention, utility model, and industrial design. Invention patents account for 29% of total Chinese applications in 2020. As applications of the last two do not require substantial review, we focus on invention patents, referred to as "patent" throughout the paper.

2011, China replaced the US as the world's No.1 patent application country. By 2020, this share increased to 45.69% of the world's total. The number of patents per researcher in China started at a much lower level in the early 1990s. Both countries progressed at a comparable rate in the 1990s. The US-China gap shrank in the 2000s, suggesting China's technological catch-up in that decade. Over the recent 10-15 years, when the Chinese government set quantity targets and adopted various innovation subsidies, patents per researcher in China have increased much faster than in the US. By 2018, an average Chinese researcher produced almost twice as many patents as their US counterparts.[4]

Figure 2.1: Patent Quantity in China and the United States



(a) No. of Patent Applications (mn) (b) Patents per Researcher

Note: This figure shows the number of patents (millions) and patents per researcher in China and the US.

The patent surge raises concerns on whether Chinese innovators are becoming more productive or are incentivized to focus primarily on quantity while ignoring the underlying quality of patents, which we, by assembling a panel data of Chinese innovating firms, turn next to.

2.2 Firm-Level Data and Patent Quality

Data Source. We construct an input-output panel data of firm-level R&D activities from three sources: (i) Annual Survey of Industrial Enterprises (ASIE), covering above-scale

[4] Appendix A.1 confirms that the evolution of patent grants exhibits very similar patterns.

Chinese industrial firms from 1998-2013; (ii) a supplementary Firm Innovation Activity Database containing industrial firms' R&D personnel and expenditures from 2008-2014; (iii) Innography and Orbis Patent Database, providing patent information from 1985 onwards. We focus on applied and eventually granted patents and restrict to domestic firms with records of at least one invention patent during the sample period.[5]

We label a Chinese patent as radical if it is cited by at least one US patent and the gap in application years between the cited and citing patents must be within 5 years, and as incremental otherwise.[6] Figure 2.2 presents the evolution of patent quality in our firm-level sample. The left panel shows the share of radical patents, which displays a clear rising trend from 1998 to 2007 — which partly reflects increasing international exposure and the associated learning process of Chinese innovating firms starting from a low level (Baslandze et al., 2021) — and declines from 2008 onward. The post-2008 decline is more significant if compared to the pre-2008 trend. In the right panel, we show the relative quality of incremental to radical patents, defined as the ratio of the average number of forward citations of the former to the latter. That ratio is relatively stable from 1998 to 2007, but steadily declines after 2008. As more incremental patents are created, their average quality declines, suggesting a crowding effect of incremental innovations.[7]

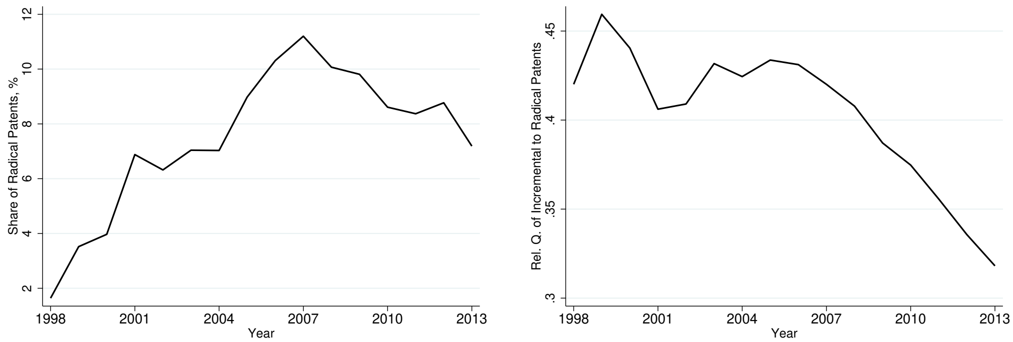
Lastly, we present the firm-level skill composition of researchers, which is later used to infer the heterogeneous innovation skill intensities in model calibration. We identify a firm as high-type if it creates at least one radical patent from 2008 to 2013, and as low-type otherwise. Among an innovating firm's R&D personnel, we label those with a medium or

[5]Appendix A.2.1 describes data sources. Construction process of the firm-level sample is given in Appendix A.2.2. Appendix A.2.3 details variable construction.

[6]We downloaded and updated patent data in October 2022, which is 9 years after the last year (2013) in our sample. The five-year restriction is to minimize the impact of truncation.

[7]These findings are robust to alternative definitions of radical patent and are evident across firm demographic characteristics, including ownership, exporting status, industries, patent categories, entrants vs. incumbents, internal vs. external patents as detailed in Appendix A.2.5.

Figure 2.2: Evolution of Patent Quality in China



(a) Share of Radical Patents (%)

(b) Relative Quality of Incremental to Radical Patents

Note: A patent is classified as radical if it is cited by 1 U.S. patent and the gap in application years between the cited and citing patents must be within 5 years. The relative quality of incremental to radical patents is the ratio of the average number of forward citations of the former to the latter. The high volatility at the beginning few years is largely due to limited sample size, as shown in Table A.3.

senior professional title as skilled personnel. Skill intensity, defined as the ratio between skilled personnel and total R&D personnel, is 34.12% for high-type firms and 25.42% for low-type firms in the 2011-2013 period. More details are provided in Appendix A.2.6.

Facing quantity-based subsidies, firms may find it profitable to maximize the number of innovations and shift from radical innovations to cheaper incremental ones with lower quality. A micro-level quantity-quality trade-off that innovating firms experience and quantity targets may help explain the observed patent surge along with their quality decline. In that spirit, we build a Schumpeterian growth model featuring heterogeneous innovations to evaluate the aggregate impact of quantity-based innovation policies.

3 The Model

This section develops a growth model with heterogeneous innovations to study the economic consequences of quantity-based subsidies. The model is in continuous time, denoted by t. The economy admits a representative household that maximizes the discounted sum of utility

$$U = \int_0^\infty \exp(-\rho t)\, \frac{C(t)^{1-\nu} - 1}{1-\nu}\, dt, \tag{1}$$

where $\rho > 0$ is the discount factor, ν the elasticity of intertemporal substitution, and $C(t)$ is the flow of final good consumed.

There is a final good produced competitively by packaging a continuum of intermediate varieties

$$Y(t) = \frac{1}{1-\epsilon} N(t)^\epsilon \int_0^1 q_\omega(t)^\epsilon y_\omega(t)^{1-\epsilon} d\omega, \tag{2}$$

where $y_\omega(t)$ is the quantity of intermediate good ω, and $q_\omega(t)$ denotes its quality. $\epsilon \in (0,1)$ governs the value-added share of intermediate varieties. $N(t)$ is the number of packagers whose total supply is fixed at 1. The final good producers' demand for intermediate variety ω is given by $p_\omega = q_\omega^\epsilon y_\omega^{-\epsilon}$.

Production. Each intermediate good $\omega \in [0,1]$ is produced by a firm that currently owns the leading technology in that product line, that is, offering the highest quality q_ω.[8] Denote Ω_f the set of product lines owned by an individual firm f, and $Q_f \equiv \{q_\omega, \omega \in \Omega_f\}$ its product portfolio. Denote n_f as the cardinality of the set Q_f, which represents the number of product lines that the firm owns and we refer to as "firm size". For simplicity, we drop the subscript f when it causes no confusion. A firm that loses all product lines exits the economy permanently, so we have $n \geq 1$ for incumbent firms.

Production of intermediate goods uses unskilled labor $y_\omega(t) = \bar{q}(t)\ell_\omega(t)$, where $\bar{q}(t) \equiv \int_0^1 q_\omega(t) d\omega$ is the economy-wide average productivity at time t, capturing a cross-firm spillover of innovations. We follow the standard approach in the Schumpeterian growth

[8] We use "intermediate good", "intermediate variety" and "product line" interchangeably in the paper.

literature and assume a two-stage price-bidding game (Acemoglu et al., 2012).[9] In equilibrium, the firm owning the leading technology can charge a monopolistic price until being replaced by a successful innovator. Given this setting, the firm that owns the leading technology in product line ω will charge a constant markup, and the profit is $\pi_{\omega} = \pi q_{\omega}$, where $\pi \equiv \epsilon \left[(1-\epsilon)\bar{q}/w^{\ell}\right]^{\frac{1-\epsilon}{\epsilon}}$ is the economy-wide average profit level and a constant on the balanced growth path as shown below.

R&D Heterogeneity. Intermediate goods firms also spend on R&D to pursue innovations. R&D efforts are assumed to be undirected. Upon a successful innovation, a firm improves the quality of a random product line by a step-size from its current frontier.[10] Innovations are heterogeneous; specifically, two kinds of innovations exist: radical (d) vs. incremental (m). We use subscript i to index innovations, i.e., $i = d, m$. Incremental innovations build on one existing radical innovation. The quality improvement associated with radical innovations is fixed and large, while that of incremental innovations is small and gradually diminishes toward zero.

R&D uses skilled labor, unskilled labor, and research time as inputs. Each firm is endowed with 1 unit of non-tradable research time. We introduce research time to capture R&D inputs that are scarce and non-tradable, such as a manager's time to supervise projects, etc.[11] When a firm owning n product lines hires h_i units of skilled labor, ℓ_i units of unskilled labor, and uses e_i fraction of its research time to pursue innovations of kind i, it adds one more product line to its portfolio at the following Poisson flow rate

[9] In stage 1, firms decide whether to pay an arbitrarily small but positive market-entry cost. In stage 2, all firms that have paid the cost in stage 1 compete in a Bertrand competition. The firm that owns the leading technology and produces the highest quality goods would announce a limit price, which makes all others earn a non-positive profit in stage 2. Therefore, they optimally decide not to enter and compete in stage 1.

[10] We focus on external innovations, which are the vast majority among innovations in China, as shown in Table A.9. We also craft an extension allowing for internal innovations in Appendix D.4.

[11] The scarce R&D inputs induce a firm-level trade-off between radical and incremental innovations, which becomes more clear once we present the innovation cost function.

$$X_i = z_i\, n^{1-\phi}\left(e_i\, h_i^{\gamma_i} \ell_i^{1-\gamma_i}\right)^{\phi}, \tag{3}$$

where $z_i \geq 0$ is the firm's R&D productivity in pursuing kind i innovations. Parameter $\phi \in (0, 1)$ is the elasticity of successful innovations to R&D expenditures. Parameter $\gamma_i \in (0, 1)$ governs the skill intensity of kind i innovations. Following empirical results in Section 2.2, we assume $\gamma_m < \gamma_d$; that is, incremental innovations are less skill-intensive than radical ones.

If firm f successfully adds a product line ω to its portfolio following a radical innovation, it raises the quality of product ω by

$$q_\omega(t_+) = q_\omega(t) + \lambda \bar{q}(t), \tag{4}$$

where $\lambda > 0$ is an exogenous parameter governing the step-size of radical innovations.

The quality improvement following a successful incremental innovation, however, depends on its distance from the most recent radical innovation, i.e., times of incremental improvements already created in that product line. Denote τ_ω this distance for product line ω, that is, if product line ω is experiencing the τ_ω-th incremental innovation from its most recent radical one, the quality improvement is

$$q_\omega(t_+) = q_\omega(t) + \eta \alpha^{\tau_\omega - 1} \bar{q}(t), \tag{5}$$

where $\eta \in (0, \lambda)$ governs the initial step-size, and $\alpha \in (0, 1)$ governs how fast the effect diminishes. The effect of incremental innovations weakens until a radical one arrives and resets the clock. This setting captures a positive externality of radical innovations, and a negative externality of incremental ones, with the latter helping reconcile the crowding effect documented in Section 2.2.

Now we are ready to derive the associated R&D cost functions. Following Klette and Kortum (2004), it is useful to transform variables into their *per line* correspondences. Denote $x_d \equiv X_d/n$ as the radical *innovation intensity* per line, $x_m \equiv X_m/n$ the incremental *innovation intensity* per line, and w^h and w^ℓ competitive wage rates for skilled and unskilled labor, respectively. For an individual firm whose innovation intensities are x_d and x_m, the associated function of R&D cost *per line* is given by[12]

$$R(x_d, x_m; z_d, z_m) = \left[\Theta_d(x_d)^{\frac{1}{2}} + \Theta_m(x_m)^{\frac{1}{2}}\right]^2, \tag{6}$$

where $\Theta_i(x_i) \equiv \Delta_i \left(w^h\right)^{\gamma_i} \left(w^\ell\right)^{1-\gamma_i} (x_i/z_i)^{\frac{1}{\phi}}$ and $\Delta_i \equiv \gamma_i^{-\gamma_i}(1-\gamma_i)^{\gamma_i - 1}$, for $i = d, m$. We further assume $\phi < 0.5$ to maintain a decreasing return to scale in R&D inputs and avoid corner solutions in x_d or x_m. The cost function indicates a clear trade-off between radical and incremental innovations at the firm level, and is a direct consequence of the scarce research time, e, required in producing both kinds of innovations.[13]

Quantity-Based Subsidy. Though innovations are heterogeneous in their magnitude of quality improvement, a successful innovation, radical or incremental, always brings the firm a new product line. At any point in time t, we assume a successful innovation embodies a certain number of patents — radical innovations correspond to radical patents and incremental innovations to incremental patents. The total number of active patents a firm holds is therefore proportional to the number of product lines the firm controls.

We define quantity-based subsidy to innovating firms as any subsidies that reward the number of active patents a firm holds, i.e., n, disregarding the underlying quality. In

[12]Under our specification of the innovation production function, a firm's innovation cost scales up linearly with the number of product lines. We craft an extension allowing for decreasing return to scale in innovation in Appendix D.3.

[13]In Appendix D.2, we relax this assumption and examine the consequences of a weakened firm-level quantity-quality trade-off.

particular, we use the form $n \times b_n \bar{q}$ in the model, where b_n denotes the detrended amount of subsidy per patent. Conceptually, the b_n term summarizes all explicit subsidies — cash or cash-like subsidies that show up in the firm's balance sheet, and implicit ones — cheaper land cost, accessibility to loans, etc., that an innovating firm receives as long as the subsidies are quantity-based.[14]

Incumbent Firms. The economy admits two types of firms regarding R&D productivity. The high-type (H) firms are capable of pursuing both radical and incremental innovations ($z_{\text{H}d}, z_{\text{H}m} > 0$). The low-type (L) are capable of pursuing only incremental innovations ($z_{\text{L}d} = 0, z_{\text{L}m} > 0$).[15] The state variables of an incumbent firm include its type $j = \text{H}, \text{L}$; its product portfolio Q; and the economy's average productivity $\bar{q}$. Denote r the interest rate and δ the creative destruction rate. The value function for a type j firm is written as

$$
\begin{aligned}
rV_j(Q,\bar{q}) - \dot{V}_j(Q,\bar{q}) = \max_{x_{jd},x_{jm}} \sum_{q_\omega \in Q} \Big\{ & \underbrace{\pi q_\omega}_{\text{profit}} + \underbrace{\delta\big[V_j(Q\backslash\{q_\omega\},\bar{q}) - V_j(Q,\bar{q})\big]}_{\text{loss from creative destruction}} \Big\} \\
& + \underbrace{n \times x_{jd}\Big[\mathbb{E}_{\omega'} V_j\big(Q \cup \{q_{\omega'} + \lambda\bar{q}\},\bar{q}\big) - V_j(Q,\bar{q})\Big]}_{\text{return from radical innovations}} \\
& + \underbrace{n \times x_{jm}\Big[\mathbb{E}_{\omega'} V_j\big(Q \cup \{q_{\omega'} + \eta\alpha^{\tau_{\omega'}-1}\bar{q}\},\bar{q}\big) - V_j(Q,\bar{q})\Big]}_{\text{return from incremental innovations}} \\
& - \underbrace{n \times R(x_{jd}, x_{jm}; z_{jd}, z_{jm})}_{\text{R\&D cost}} + \underbrace{n \times b_n\bar{q}}_{\text{quantity-based subsidy}} .
\end{aligned}
\tag{7}
$$

The first line is each product line's profit flow, plus the value loss from creative destruction. $Q\backslash\{q_\omega\}$ denotes the remaining portfolio after losing line ω to a successful innovator. The second line is the value change from a successful radical innovation of the firm, which adds product line ω' into the portfolio. The expectation is over ω' as which line the innovation lands on is random. The third line is the value change following a successful

[14]In the model, subsidizing the patent stock: $n \times b_n\bar{q}$, or subsidizing new patents: $nx \times b_x\bar{q}$, generate identical outcomes. See more details in Appendix B.4.

[15]Note $z_{\text{L}m}$ might differ from $z_{\text{H}m}$. We introduce this firm heterogeneity for quantitative purposes. The setting helps us infer R&D input structure from observable firm-level data in the quantified model.

incremental innovation. The last line includes the R&D cost and quantity-based subsidies.

The value functions are linear in the economy's average productivity $\bar{q}$, making detrending all values by $\bar{q}$ straightforward. Another useful property is that firms of the same type always choose the same innovation intensity per line, regardless of their differences in product portfolio or size.[16] By construction, low-type firms optimally choose $x_{\mathrm{L}d} = 0$. We end up tracking three innovation intensities: $x_{\mathrm{H}d}$, $x_{\mathrm{H}m}$ for high-type firms, and $x_{\mathrm{L}m}$ for low-type firms. For the remaining theoretical analysis, we focus on high-type firms, as they are the ones facing a quantity-quality trade-off between radical and incremental innovations.

Entrant Firms. At any point, there is a total mass of 1 of potential entrants pursuing incremental innovations at a fixed Poisson rate x_{E}. Upon a successful innovation, the potential entrant enters the economy with one product line in its portfolio. Entrants are of low-type by default; however, after making a one-time overhead investment of $K(p) = [-\ln(1-p) - p]\,\chi\bar{q}$, they receive probability p of becoming high-type. The value function for a potential entrant is

$$rV_{\mathrm{E}} = x_{\mathrm{E}} \left[\max_{p} \left\{ pV_{\mathrm{H}} + (1-p)V_{\mathrm{L}} - K(p) \right\} - V_{\mathrm{E}} \right], \tag{8}$$

where $V_j \equiv \mathbb{E}_{\omega'} V_j\Big(\big\{ q_{\omega'} + \eta\alpha^{\tau_{\omega'}-1}\bar{q} \big\}, \bar{q} \Big)$, $j = \mathrm{H}, \mathrm{L}$ are the expected values of a type j firm with one product line. Since entrant firms are ex-ante identical, they end up choosing the same amount of overhead investment. We denote p^* as the associated probability.

Education. The representative household also supplies a mass L of workers, each facing a constant death rate of $d > 0$. At each point, a flow dL of young workers join the

[16]This property is obtained as the elasticity of innovation on firm size n is set as $1-\phi$ in the innovation production function, so the innovation cost scales up linearly with the number of product lines. We craft an extension allowing for decreasing return to scale in innovation in Appendix D.3.

economy, who work as unskilled without any investment in education; however, they can spend time in school to become skilled. Upon entry, each individual randomly draws a talent type θ from a *Pareto* distribution.[17] It requires $1/\theta$ units of education service for an individual of type θ to become skilled. Education service is produced by existing skilled labor employed in education at the competitive wage rate and with technology $S = \xi\, h^{\text{teacher}}$, where $\xi > 0$ captures the overall efficiency of the economy's education infrastructures. Getting education is a preferable choice if and only if the expected lifetime return from doing so — earning a skilled wage minus paying the education cost — exceeds the lifetime value of earning an unskilled wage.[18]

3.1 Equilibrium

We focus on a balanced growth path equilibrium, in which the average productivity of the economy, $\bar{q}(t)$, grows at a constant rate g, while other aggregate variables grow proportionally and all relevant distributions are stationary. We track two distributions, one over τ, i.e., the step-size of incremental innovations, and the other over n, i.e., firm sizes.[19] Denote δ_d and δ_m the creative destruction rate due to radical and incremental innovations, respectively. The expected step-size of an incremental innovation is given by

$$\bar{\eta} = \eta \Big/ \left(\alpha + \frac{1-\alpha}{\delta_d/\delta} \right). \tag{9}$$

The expected quality improvement from incremental innovations decreases with a faster decay rate, i.e., a smaller α. Additionally, as the fraction of radical innovations in the economy, δ_d/δ, decreases, the expected step-size also becomes smaller. This property allows us to simultaneously explain the decline in the share of radical patents, as well as the widening gap between incremental and radical patents documented in Figure 2.2.

[17]More specifically, we use Pareto distribution $\mathbb{P}\left\{\theta \le \tilde{\theta}\right\} = 1 - \tilde{\theta}^{-2}$, for $\tilde{\theta} \in [1, \infty)$.

[18]Appendix B.1 shows that young people obtain education when surpassing a certain talent threshold θ^*, the value of which depends on the skill premium as well as productivity in the education sector.

[19]Detailed derivations regarding these two stationary distributions are given in Appendix B.2.

Denote $\mu_{j,n}$ the mass of type j firms who own n product lines. We can write the creative destruction rates $\delta_d = \sum_n \mu_{\mathrm{H},n} \times n x_{\mathrm{H}d}$; $\delta_m = \sum_j \sum_n \mu_{j,n} \times n x_{jm} + x_{\mathrm{E}}$. The aggregate rate of creative destruction is $\delta = \delta_d + \delta_m$. Moreover, the total number of active product lines sums to 1. Formally, the following proposition holds.[20]

Proposition 1. *Definition of the creative destruction rate δ guarantees that $\sum_j \sum_n \mu_{j,n} \times n = 1$.*

3.2 Properties of the Economy

The Quantity-Quality Trade-off. High-type firms in the economy face a quantity-quality trade-off between creating more innovations and creating better innovations. Innovation subsidies may impact such trade-offs in an undesired way, as seen from an individual firm's optimal decisions. For expositional convenience, we drop the firm type subscript j whenever it causes no confusion.

Innovating firms' value function takes the form $V(Q, \bar{q}) = \sum_\omega A q_\omega + nB\bar{q}$. The first term denotes profit from owning product lines, while the second contains net values from R&D. Regarding firms' choices over innovations, we have the following proposition.[21]

Proposition 2. *The ratio between radical and incremental innovation intensities satisfies*

$$\frac{x_d}{x_m} \propto \underbrace{\frac{A(1+\lambda)+B}{A(1+\bar{\eta})+B}}_{\text{innovation return}} \times \underbrace{\left(\frac{w^h}{w^\ell}\right)^{-(\gamma_d-\gamma_m)}}_{\text{input structure}}. \tag{10}$$

The term "innovation return" on the right-hand side captures the ratio of returns between radical and incremental innovations. The direct return of radical innovations is from its productivity improvement effect, as captured by $A(1+\lambda)$. Similarly, that of incremental

[20] Proof of Proposition 1 can be found in Appendix B.3.

[21] Appendix B.4 derives the value function and proves Proposition 2.

innovations is captured by $A(1+\bar{\eta})$. The fact $\lambda > \bar{\eta}$ indicates that the direct return of radical innovations is greater.

The indirect return, B, is identical for both kinds of innovations and is largely affected by subsidies. A sizable subsidy shrinks the gap in total returns between radical and incremental innovations and raises firms' incentive to pursue proportionately more of the latter. The term "input structure" indicates that, if all firms are doing more R&D, the equilibrium skill premium will rise, making skill-intensive radical innovations more expensive. Consequently, firms are further incentivized to pursue incremental innovations.

Growth and Welfare. Along a balanced growth path, the aggregate welfare is

$$U = \frac{1}{1-\nu}\left[\frac{{C_0}^{1-\nu}}{\rho-(1-\nu)g} - \frac{1}{\rho}\right]. \tag{11}$$

A critical determinant of welfare is the aggregate growth rate, $g = \delta_d\lambda + \delta_m\bar{\eta}$, where $\bar{\eta}$ denotes the expected step-size of incremental innovations. As δ_d and δ_m represent aggregate quantity of radical and incremental innovations, the growth rate can be viewed as a weighted sum of their step-sizes.[22] Accordingly, the growth rate differential, e.g., between economies with and without a particular policy, can be decomposed into

$$\Delta g = \underbrace{\Delta\delta \times \left[\frac{\delta_d}{\delta}\lambda + \left(1-\frac{\delta_d}{\delta}\right)\bar{\eta}\right]}_{\text{(i) quantity-creative destruction}} + \underbrace{\delta \times \left[\Delta\frac{\delta_d}{\delta} \times (\lambda-\bar{\eta})\right]}_{\text{(ii) quality-composition}} + \underbrace{\delta \times \left[\left(1-\frac{\delta_d}{\delta}\right) \times \Delta\bar{\eta}\right]}_{\text{(iii) quality-crowding}}. \tag{12}$$

The first term refers to the quantity channel, while the second and third are the quality channels. The *quantity-creative destruction* term captures that the aggregate growth rate changes if a policy induces changes in the aggregate creative destruction rate, δ, or

[22] Difference between the contribution of one extra radical innovation versus that of an incremental one is $(\lambda-\bar{\eta}) + \delta_m(\partial\bar{\eta}/\partial\delta_d - \partial\bar{\eta}/\partial\delta_m)$. The latter term corresponds to the externality induced by our particular way of modeling the quality decay of incremental innovations.

equivalently the total number of innovations. As aforementioned, we focus on external innovations which are the vast majority in China. A policy that changes the aggregate share of radical innovations, δ_d/δ, further affects aggregate growth through (a) it changes the composition of radical and incremental innovations, whose impacts on productivity are different, captured by the second *quality-composition* channel; and (b) changing the average number of incremental innovations following a radical one in any product line changes the average productivity impact of incremental innovations, which we label the third *quality-crowding* channel. The three channels connect to the patent surge in Figure 2.1, and the two facts regarding patent quality decline in Figure 2.2 in the Chinese context.

A quantity-based subsidy might promote overall growth and welfare through the quantity channel; however, the positive effect could be compromised or even overwhelmed if the subsidy negatively impacts innovation quality. Which effect dominates is a quantitative issue addressed in the following section.

4 Quantitative Analysis

This section first calibrates the model using Chinese data, and evaluates the impact of *quantity-based* subsidies on patent quantity surge, quality decline, and the overall TFP growth. We then analyze a planner's problem, which yields a constrained first-best, and propose an alternative, quality-biased, innovation policy — subsidizing the skill, which we show effectively recovers the planner's allocation.

4.1 Calibration and Model Fit

To calibrate the model's benchmark economy to 2011-2013 aggregate- and firm-level data, we further include two extra policy parameters: the corporate tax rate u — hence firms' profit flow changing from πq to $(1-u)\pi q$ — and the R&D tax credit multiplier b_r, i.e., the total R&D expenditures changing from $R(x_d, x_m)$ to $(1 - b_r u)R(x_d, x_m)$.

Calibration Strategy. The extended model has 20 parameters. We start with those that can be externally calibrated, directly inferred, or taken from the literature. We set the discount rate $\rho = 0.02$. For the inverse intertemporal substitution elasticity, we set $\nu = 3$ in the baseline and check the robustness with alternative values. The elasticity of substitution in final goods production ϵ is set to match a profit rate of 22% among ASIE firms. The total population L is normalized to 1.

In the R&D sector, we follow Acemoglu et al. (2018) relying on microeconometric innovation literature, and estimate an innovation elasticity parameter $\phi = 0.49$.[23] Without loss of generality, we assume that the initial step-size of incremental innovations $\eta = \alpha\lambda$, with the latter two parameters calibrated internally. In the education sector, we set the death rate d so that an individual works for around 35 years. We set u and b_r to match a 25% corporate tax rate and a 150% tax credit multiplier in China.[24]

The remaining 11 parameters are internally calibrated to moments, which, unless stated otherwise, are calculated from our firm-level sample for 2011-2013. For each of the parameters, we pick a most informative moment implied by the model. The first set of remaining parameters regards R&D productivity. The model contains three such parameters: radical and incremental innovation productivity for high-type firms, $z_{\mathrm{H}d}$ and $z_{\mathrm{H}m}$; and incremental innovation productivity for low-type firms, $z_{\mathrm{L}m}$. To discipline $z_{\mathrm{H}m}$ and $z_{\mathrm{L}m}$, we use the average R&D intensity, defined as the ratio of total R&D expenditures to value-added,[25] of high- and low-type firms. The ratio $z_{\mathrm{H}d}/z_{\mathrm{H}m}$ affects the share of radi-

[23] We regress the number of patents on (log) R&D expenditure, controlling for (log) R&D staff, as well as year, location (province), industry, ownership types, and establishment year fixed effects, using the Poisson quasi-maximum likelihood estimator with robust standard errors clustered at the firm level. The estimate is quite close to values used in the literature.

[24] The values of all externally calibrated parameters and their sources are summarized in Table C.1.

[25] In the model, firms with the same R&D productivity choose the same level of R&D expenditures per

cal innovations high-type firms choose to pursue. Therefore, we use the share of radical patents, to discipline the value of $z_{\mathrm{H}d}$.

The second set of parameters regards skill intensities in R&D: γ_d and γ_m. Recall that low-type firms are creating only incremental patents, so we can use the observed skill intensity of them to discipline γ_m. With the value of γ_m determined, we can further discipline γ_d by targeting the observed skill intensity of high-type firms.

We pin down the amount of subsidy, b_n, by the aggregate subsidy-to-R&D expenditure ratio.[26] Following the formula of aggregate growth rate, we set the step-size of radical innovations, λ, to match an annual TFP growth rate of 1.97% in 2008-2014 estimated by Bai and Zhang (2017). As we map step-size of incremental innovations to the number of forward citations received by incremental patents, the decay rate α, which determines the average step-size of incremental innovations, can be disciplined by the ratio between average forward citations received by incremental and radical patents.[27]

We use the skill premium to discipline productivity in the education sector ξ, which determines the total supply of skilled labor and the equilibrium wage rates.[28] Consistent with the model's setting, entrant firms' innovation rate x_{E} is disciplined by the share of patents created by new entrants in the economy. The cost coefficient for entrants to become high-type, χ, is set to match the percentage of high-type firms among incumbents. In the end, we jointly calibrate all 11 parameters to minimize the total sum of distance

line, but their value-added may differ due to idiosyncratic draws of product quality. "Average" means a within-type semi-aggregation that gives a "representative" value-added for each type of firm.

[26] In the data, we define subsidy as the sum of government research funds, subsidy to innovative firms in the ASIE database, and HTE tax exemptions, where HTE refers to High-Tech Enterprises. HTE recognition, also known as the InnoCom Program, is a critical pro-innovation subsidy program China has initiated. For more discussion on HTEs, please see Appendix C.4.

[27] Appendix C.2 provides more details.

[28] To obtain the skill premium, we run a Mincer regression using data from the Urban Household Survey 2009, the coefficient in front of the dummy for "graduate degree" is 2.43. Specifically, we regress wage on the education group dummy controlling for household age, age squared, gender, race, and marital status.

between model-generated and data moments. Table 4.1 summarizes the internally calibrated parameters and their corresponding target moments.

Table 4.1: Internally Calibrated Parameters

Para	**Equation**	**Meaning**	**Target**
$z_{\mathrm{H}d}$	(7)	H-type's radical productivity	share of radical innov.
$z_{\mathrm{H}m}$	(7)	H-type's incremental productivity	H-type's R&D intensity
$z_{\mathrm{L}m}$	(7)	L-type's incremental productivity	L-type's R&D intensity
γ_d	(3)	skill intensity in radical innov.	H-type's skill intensity
γ_m	(3)	skill intensity in incremental innov.	L-type's skill intensity
b_n	(7)	quantity-based subsidy	subsidy-to-R&D ratio
λ	(4)	step-size of radical innov.	TFP growth rate
α	(5)	speed of quality decay	average citation ratio
ξ		education productivity	skill premium
x_{E}	(8)	entrants' innov. rate	entrants' patent share
χ		cost of becoming H-type	fraction of H-type incumbents

Model Fit. Table 4.2 presents the calibration results and model fit, the benchmark model well replicates the targeted moments. We show how the total distance and each moment change with respect to the corresponding parameter's value in Appendix C.3.

Table 4.2: Benchmark Calibration

Para	$z_{\mathrm{H}d}$	$z_{\mathrm{H}m}$	$z_{\mathrm{L}m}$	γ_d	γ_m	b_n	λ	α	ξ	x_{E}	χ
Value	1.029	1.038	1.016	0.796	0.453	0.029	0.158	0.862	0.035	0.068	0.138
Data (%)	8.01	17.78	15.02	34.12	25.42	20.42	1.97	33.28	243	21.00	26.98
Model (%)	8.01	17.49	15.02	34.08	25.41	20.44	1.97	33.27	243	20.98	26.99

Estimates of γ_d and γ_m confirm that R&D activities pursuing radical innovations rely more heavily on skilled labor than incremental ones. The relatively large difference between the two values, 0.796 vs. 0.453, is necessary to account for the observed 9% gap between skill intensities of high- and low-type firms, as more than 70% of innovations created by high-type firms, are incremental.

To give a sense of the magnitude of quantity-based subsidy b_n, we contrast it to the average value-added of innovating firms, and the ratio is slightly above 3%. Our estimate of λ implies that a radical innovation improves the quality of a product by 15.8%. This number is close to that obtained in the literature. For example, Akcigit and Kerr (2018) estimated a step-size of 11.2%, while Acemoglu et al. (2018) reported a step-size of 13.2%. Our estimate of $\alpha = 0.862$ is lower than what is reported in Akcigit and Kerr (2018) about US patents, implying a faster quality decay among incremental patents in China.

The calibrated model generates a creative destruction rate $\delta = 32.31\%$. As our paper corresponds innovation to patents, we define a patent-level creative destruction rate in a year as the ratio of newly created patents to that of the patent stock. For the 2011-2013 sample period, we estimate a patent-level creative destruction rate that ranges from 30% to 34%, which is consistent with the model counterpart and close to other estimates in the literature, e.g., Branstetter et al. (2023).

The model also predicts that firms with a higher innovation intensity have a larger expected size. We examine the model-generated relative size between high- and low-type firms and contrast it to the data counterpart (Appendix Table C.2). Lastly, we compare the patent number distribution in the model and the data (Appendix Figure C.3), and provide supportive evidence that the magnitude of quantity-quality trade-off implied by the calibrated model is in line with what's revealed in the data. Details of the aforementioned model fit examinations can be found in Appendix C.4.

4.2 Effects of Quantity-Based Subsidies

We are now ready to evaluate the impact of quantity-based subsidies. To that end, we compare the baseline outcome with a counterfactual economy in which all such subsidies are shut down, i.e., $b_n = 0$. Table 4.3 presents the results.

Table 4.3: Innovation Quantity & Quality in the Baseline (B.M.) and Counterfactual Economy (C.F.) w./o. Quantity-Based Subsidies

Variable	Meaning	B.M.	C.F.	Δ_{Model}	Δ_{Data}	$\frac{\Delta_{\text{Model}}}{\Delta_{\text{Data}}}$
$\delta - x_{\text{E}}$	incumbent innovation	25.53%	23.18%	10.14%	34.57%	29.33%
δ_d / δ	radical share	8.01%	10.39%	-22.91%	-40.89%	56.03%
$\bar{\eta} / \lambda$	step-size ratio	33.27%	39.27%	-15.28%	-20.27%	75.38%

Note: Δ_{Model} represents changes from the counterfactual to the model benchmark, Δ_{Data} is changes between the pre- and post-2008 period, both columns are presented in relative terms. Step-size ratio denotes the relative step-size of incremental to radical innovations.

Quantity-based subsidies implemented in China generate a relative increase of 10.14% in innovation quantity, measured by incumbents' creative destruction. However, such subsidies reduce overall innovation quality in two dimensions. Firstly, the share of radical innovations declines from 10.39% to 8.01%, a relative decrease of 22.91%. Secondly, the relative quality of incremental to radical innovations decreases, indicated by a decline in the average step-size ratio from 39.27% to 33.27%, or a relative decrease of 15.28%.

In the data, compared to the pre-2008 period, we see a relative increase of 34.57% in patent quantity, a relative decrease of 40.89% in the share of radical patents, and a relative decrease of 20.27% in the average citation ratio between incremental and radical patents.[29] Our quantified model implies that quantity-based subsidies account for about 29% of the patent surge, 56% of the radical share decline, and 75% of the widening gap between incremental and radical innovations, observed in the post-2008 period.

By changing firms' innovation incentives, quantity-based subsidies further affect aggregate growth and welfare. In our exercise, the aggregate growth rate decreases from 2.16% in the counterfactual economy without subsidies to 1.97% in the baseline. Bai and Zhang (2017) report that the Chinese TFP growth rate decreases by 1.91 percentage points, from

[29] See Appendix D.1 for details of the estimation. Essentially we fit a linear pre-2008 trend and extrapolate that trend to obtain the "natural" level for years after 2008. By calculating the deviation of actual values from these predicted values in relative terms, we obtain and report the relative changes here.

3.88% in 2001-2007 to 1.97% in 2008-2014. A drop of 0.19 percentage points in the model's TFP growth rate accounts for about 10% of the change between the pre- and post-2008 periods in the data.

We then follow equation (12) and decompose such effects into three channels, as shown in Table 4.4. Subsidies raise innovation quantity, leading to a 0.17 percentage points increase in aggregate growth rate; however, this positive quantity effect is overwhelmed by the negative quality effects. A pool of proportionately less radical innovations reduces growth. This channel, which corresponds to δ_d/δ, brings a 0.07 percentage points drop in aggregate growth. In addition, the average productivity enhancement from incremental innovations falls as quantity-based subsidies induce more incremental R&D trials. This last channel, which corresponds to $\bar{\eta}/\lambda$, brings a 0.26 percentage points drop in aggregate growth rate. Overall, the quality-crowding effect dominates, generating a negative net effect on growth. As a result, the introduction of quantity-based subsidies causes a welfare loss of 3.31% to the economy.

Table 4.4: Growth Decomposition

Δ_{Growth}	(i) quantity	(ii) quality-composition	(iii) quality-crowding	
-0.19	0.17	-0.07	-0.26	(p.p.)
	-89.47%	36.84%	136.84%	

Note: For each of the channels, we add the corresponding change in (i) δ; (ii) δ_d/δ; (iii) $\bar{\eta}$ to the pre-2008 economy, and see how it affects the aggregate growth rate. The second row shows the contribution of each channel, calculated by dividing the corresponding number by -0.19 p.p.

To summarize, although quantity-based subsidies promote overall innovations, they also skew R&D efforts toward incremental innovations, hence imposing negative effects on growth and welfare. Among the undesirable consequences, quality-crowding, i.e., the average quality of incremental innovation declines as more such innovations are pursued, accounts for most of the losses. Note the growth and welfare implications are based on the comparison between two balanced growth paths, while what happens on the transitional path is not addressed in the analysis. Last, the assumption of scarce research time, together with parameter values regarding the degree of quality decay and heterogeneity in skill intensities, determines the strength of the three channels and the sign of the net impact. We present a formal discussion of them in Appendix D.2.

The baseline model, designed to deliver a clean characterization of the firm-level quantity-quality trade-off and facilitate aggregation, is subject to two oversimplification concerns. One is that the innovation cost scales up linearly with firm size, the other resides in the assumption that all innovations trigger creative destruction, i.e., being external. To alleviate these concerns, we first extend the model to allow for decreasing return to scale in pursuing innovations. In another extension, we introduce internal innovations as a third choice for incumbent firms. Appendix D.3 and D.4 provide details on the model setting, calibration, and the growth and welfare implications of these two extensions, respectively. The extended models deliver similar results to those in the baseline case, reaffirming the robustness of the conclusions.

4.3 Quality-Biased Policy

This section analyzes the implementation of a constrained planner's allocation, and accordingly, proposes innovation policies that are quality-biased. In particular, we allow the planner to decide the skill supply but let individual firms produce and price as in the market economy, as we are not interested in alleviating the monopoly distortion. Since the economy contains an education sector with endogenous creation of skilled labor, the planner needs to choose a talent threshold above which the young shall obtain education, θ^*_{SP}, to maximize social welfare.

Social welfare, as defined in equation (11), is a hump-shaped function of the education threshold. An increase in skill supply initially raises welfare as it promotes innovation and growth but reduces welfare after passing a threshold, as the negative effect from a shrinking unskilled workforce and lower initial consumption level eventually dominates.[30] The socially optimal allocation of skilled labor supply is 14.36% of the population, more than tripled comparing to the market equilibrium level of 4.04%. As more skilled labor promotes innovation, the aggregate growth rate increases from 1.97% to 5.51%, and the aggregate welfare improves dramatically by 16.85%.

A natural follow-up question is whether policymakers can find implementable subsidies to recover the planner's allocation and improve welfare. Here we propose a quality-biased policy: subsidizing the skill. In particular, we consider two forms of skill subsidies. One is an *education subsidy*, with which the government covers $b_e \in [0, 100\%]$ portion of the education cost. The other is a *skilled labor subsidy*, with which the government covers $b_h \in [0, 100\%]$ portion of the skilled wage cost to innovating firms. Recall that young workers choose whether to obtain skill based on the education cost and the skill premium. Therefore, policymakers are able to implement any desired labor supply allocation with proper combinations of b_e and b_h. If the desired skill supply is high, simultaneous usage of both subsidies is required. We formalize the argument into the following proposition.[31]

Proposition 3. *For any given $\theta^*_{SP} \in [1, \theta^*_{CE})$, there exists a set $\mathcal{G}$ of different combinations of $(b_e, b_h) \in [0, 1]^2$ to implement the allocation in a market equilibrium. Moreover,*

(i) to implement the given θ^*_{SP}, policymakers face a linear trade-off between b_e and b_h;

(ii) when the talent threshold θ^*_{SP} is low enough, the set $\mathcal{G}$ does not contain $(0, b_h)$ or $(b_e, 0)$.

Lastly, we contrast the effects of "quality-biased" skill subsidies to quantity-based subsidies in Appendix D.7.[32] Different from quantity-based subsidies, the skill subsidies help

[30] Aggregate welfare is affected by intertemporal elasticity of substitution, ν. Appendix D.5 details how we solve the planner's problem, and displays the hump-shaped welfare curves under different values of ν.

[31] See proof of Proposition 3 in Appendix D.6.

[32] We also evaluate the effects of R&D tax credit, which turn out to be similar to quantity-based subsidies in our framework.

improve both innovation quantity and quality, and promote growth and welfare, as they raise the skill supply and reduce equilibrium skill premium, therefore are biased toward more skill-intensive radical innovations.

5 Conclusion

Motivated by the Chinese patent quantity surge and quality decline, we construct a Schumpeterian growth model featuring heterogeneous innovations to study the impact on growth and welfare of quantity-based subsidies widely adopted in China since the middle 2000s. We decompose the impact into positive quantity and negative quality channels. The model-based quantitative analysis shows that such subsidies reduce the aggregate welfare by 3.31%, as the negative *quality-crowding* channel dominates. We further evaluate welfare gains under a constrained planner's problem, and propose quality-biased skill subsidies which effectively recover the optimal allocation.

We necessarily abstract from other essential features, e.g., transition from imitation to innovation, to focus on the quantity-quality trade-off. In addition, the way skill accumulation is modeled is simplified to keep the framework tractable. In reality, patent subsidies work immediately, while building up a skill pool takes generations of time. Skill subsidies can also take many forms in the real world, for example, attracting overseas-trained talents to work at home seems an important channel for China's technology catch-up. We leave detailed investigations along these dimensions for future research.

References

Acemoglu, Daron, Gino Gancia, and Fabrizio Zilibotti, "Competing Engines of Growth: Innovation and Standardization," *Journal of Economic Theory*, 2012, *147* (2), 570–601.

_ , Ufuk Akcigit, and Murat Alp Celik, "Radical and Incremental Innovation: The Roles of Firms, Managers and Innovators," *American Economic Journal: Macroeconomics*, 2022, *14* (3), 199–249.

_ , _ , Harun Alp, Nicholas Bloom, and William Kerr, "Innovation, Reallocation, and Growth," *American Economic Review*, 2018, *108* (11), 3450–91.

Akcigit, Ufuk and William R Kerr, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

_ , Jeremy G. Pearce, and Marta Prato, "Tapping in Talent: Coupling Education and Innovation Policies for Economic Growth," *NBER Working Paper*, 2020.

Ang, James S, Yingmei Cheng, and Chaopeng Wu, "Does Enforcement of Intellectual Property Rights Matter in China? Evidence from Financing and Investment Choices in the High-Tech Industry," *Review of Economics and Statistics*, 2014, *96* (2), 332–348.

Ates, Sina T. and Felipe E. Saffie, "Fewer but Better: Sudden Stops, Firm Entry, and Financial Selection," *American Economic Journal: Macroeconomics*, 2021, *13* (3), 304–356.

Bai, Chong-En and Qiong Zhang, "Is the People's Republic of China's Current Slowdown a Cyclical Downturn or a Long-Term Trend? A Productivity-Based Analysis," *Journal of the Asia Pacific Economy*, 2017, *22* (1), 29–46.

Baslandze, Salome, Pengfei Han, and Felipe E. Saffie, "Imitation, Innovation, and Technological Complexity: Foreign Knowledge Spillovers in China," *Working Paper*, 2021.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 2012, *97* (2), 339–351.

Branstetter, Lee G, Douglas Hanley, and Huiyan Zhang, "Unleashing the Dragon: The Case for Patent Reform in China," *Working Paper*, 2023.

Chen, Hong, Hanbing Fan, Takeo Hoshi, and Dezhuang Hu, "Do Innovation Subsidies Make Chinese Firms More Innovative? Evidence from the China Employer Employee Survey," *NBER Working Paper*, 2019.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu, "Notching R&D Investment with Corporate Income Tax Cuts in China," *American Economic Review*, 2021, *111* (7), 2065–2100.

Ding, Xuedong and Jun Li, *Incentives for Innovation in China: Building An Innovative Economy*, Routledge, 2015.

Fang, Lily H, Josh Lerner, and Chaopeng Wu, "Intellectual Property Rights Protection, Ownership, and Innovation: Evidence from China," *The Review of Financial Studies*, 2017, *30* (7), 2446–2477.

Haley, Usha C. V. and George T. Haley, *Subsidies to Chinese Industry: State Capitalism, Business Strategy, and Trade Policy*, Oxford University Press, 2013.

Hu, Albert Guangzhou and Gary H Jefferson, "A Great Wall of Patents: What is behind China's Recent Patent Explosion?," *Journal of Development Economics*, 2009, *90* (1), 57–68.

Klette, Tor Jakob and Samuel Kortum, "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

König, Michael, Zheng Michael Song, Kjetil Storesletten, and Fabrizio Zilibotti, "From Imitation to Innovation: Where is All that Chinese R&D Going?," *Econometrica*, 2022, *90* (4), 1615–1654.

Li, Xibao, "Behind the Recent Surge of Chinese Patenting: An Institutional View," *Research Policy*, 2012, *41* (1), 236–249.

Nelson, Richard R. and Edmund S. Phelps, "Investment in Humans, Technology Diffusion, and Economic Growth," *The American Economic Review*, 1966, *56* (1/2), 69–75.

Vandenbussche, Jerome, Philippe Aghion, and Costas Meghir, "Growth, Distance to Frontier and Composition of Human Capital," *Journal of Economic Growth*, 2006, *11* (2), 97–127.

Wei, Shang-Jin, Jianhuan Xu, Ge Yin, and Xiaobo Zhang, "Mild Government Failure," *NBER Working Paper*, 2023.

Online Appendix of "Haste Makes Waste? Quantity-Based Subsidies under Heterogeneous Innovations" (Not for Publication)

A Appendix: Data and Facts

A.1 Institutional Background and Patent Quantity

Table A.1 lists quantity targets set by Chinese central and local governments in the late 2000s and early 2010s.

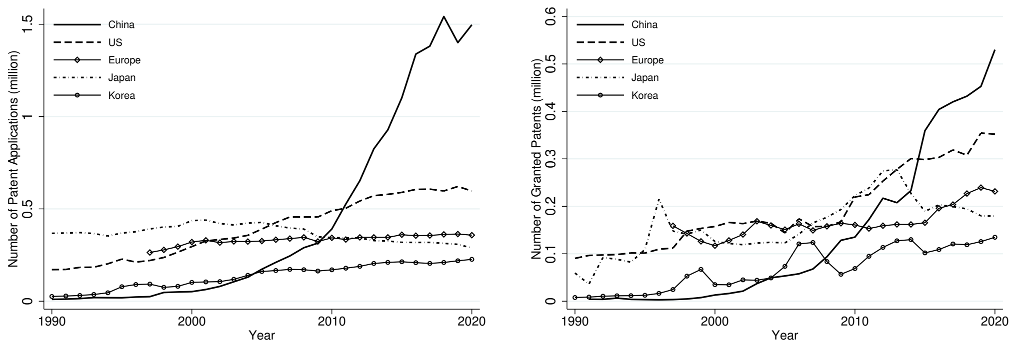
Table A.1: Quantity Targets set by the Central and Selected Local Governments

Policy Year	Target Period	Quantity Target
Central Government		
2010	2011-20	Patents to reach 2 mil. & rank Top 2 in the world in 2015 Patents per 1 mil. pop. to increase by 100% by 2015 and 300% by 2020
Beijing City		
2010	2011-15	Patent applications (*resp.* grants) per 10,000 pop. to reach 20 (*resp.* 8) by 2015
2015	2016-20	Patents per 10,000 pop. to reach 80 by 2020
Shanghai City		
2010	2011-20	Patent grants per 1 mil. pop. to reach 600, and patents per 10,000 pop. to reach 30, in 2015; both criteria to double in 2020
Guangdong Province		
2007	2007-20	Patent applications per 1 mil. pop. to reach 200 in 2010 and to increase more than 15% annually
Heilongjiang Province		
2011	2011-20	Patents per 10,000 pop. to surpass 2.1 by 2015
Guizhou Province		
2017	2016-20	Patents per 10,000 pop. to reach 2.5 by 2020

Data Source: The national targets are from National Patent Development Strategy 2011-2020. Local targets are from local Intellectual Property Development Strategy or Five Year Plans.

Figure A.1 presents the number of applied & granted patents in China and other major patent-holding countries. Table A.2 presents the number of researchers per million inhabitants in China and selected countries. Figure A.2 shows the number of patent applications per researcher in China, the US, and G5 countries. Lastly, Figure A.3 shows the patent grant rate and number of patents that have been eventually granted per researcher.

Figure A.1: Number of Applied and Granted Patents in China and Advanced Economies



(a) Patent Application (mn) (b) Granted Patents (mn)

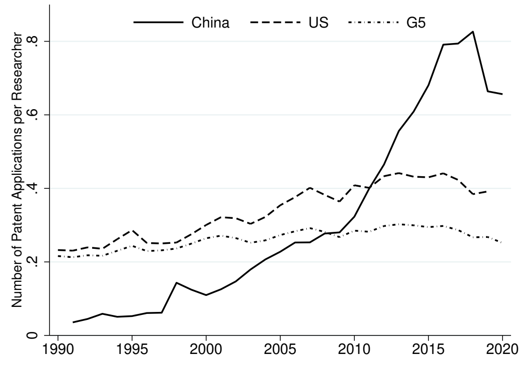
Note: This figure shows the number of applied and granted patents in China and other major patent-holding economies. The data source is World Intellectual Property Office (WIPO) IP Statistics Data Center.

Table A.2: Researchers per million Inhabitants, 2013

	China	US	Europe	Japan	France	Germany
(1)	1071.1	3984.4	2941.9	5194.8	4124.6	4355.4
(2)	0.2%	1.5%	1.8%	1.2%	1.7%	2.7%

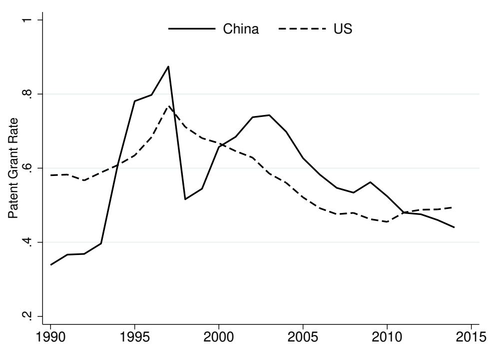
Note: Row (1) shows full-time equivalent researchers per million Inhabitants in 2013, and row (2) the share of Ph.D. degree holders in the labor force. Data source: USESCO.ORG.

Figure A.2: Number of Patent Applications per Researcher in China, US, and G5

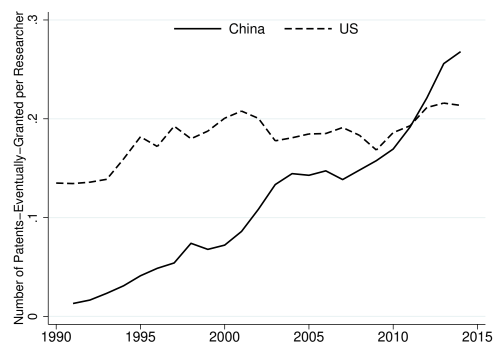


Note: Data source for No. of patents is WIPO, and for No. of full-time equivalent researchers is from OECD.Stat. This figure shows the evolution of patent applications per researcher over time. G5 includes the US, the UK, France, Germany, and Canada. It does not contain Japan or Italy as data on No. of researchers for these two countries in the OECD.Stat database are under different definitions.

Figure A.3: Patent Grant Rate and Number of Patents-Eventually-Granted Per Researcher



(a) Patent Grant Rate



(b) P-E-G Per Researcher

Note: This figure shows the patent grant rate, which is the fraction of applied patents in a given year that are eventually granted before Oct. 2020 (panel (a)), and patents-eventually-granted per researcher, which is the number of patent applications that are eventually granted per researcher (panel (b)). No. of patents that are eventually granted is calculated from patent-level data from *Innography* and *PatentsView* databases. To avoid the truncation issue, the figure only shows patents that were applied in or before 2014.

A.2 Firm-Level Sample: Data Source, Construction, and Results

A.2.1 Data Source

Annual Survey of Industrial Enterprises (ASIE). Annual Survey of Industrial Enterprises (ASIE), conducted by the National Bureau of Statistics of China (NBS), contains financial information for all state-owned enterprises, and private firms with sales above 5 million RMB before 2011 and 20 million RMB since 2011 in the industrial sector (also referred to as the "above scale" industrial firms) for the periods 1998-2013.

Innography. Innography Patent Database covers information on over one hundred million patents from various countries. In this paper, we restrict attention to patents that are applied and eventually granted in China. If a patent is filed in China in year t and eventually granted in year $t+1$, it consists of our sample of newly applied Chinese patents in year t. We supplement the Innography database with patent data from Orbis Intellectual Property.

Firm Innovation Activity Database. Firm Innovation Activity Database contains information on innovation costs and R&D expenditures, and tax cuts from the Recognition of High-Tech Enterprises, for all industrial firms that have innovation activities for the 2008-2013 period. There are in total 394,381 observations within the seven-year time period, covering approximately 120,000 unique firms.

A.2.2 Sample Construction

To construct the firm-level sample, we merge different data sources to ASIE manufacturing firms. The sample construction process consists of the following three major steps.

Step 1: Construct the 1998-2013 ASIE Sample. We follow Brandt et al. (2012) to create an unbalanced panel of firms between 1998 and 2013. We restrict the ASIE sample to the manufacturing industries, that all 4-digit CIC codes between 1300 and 4400. We drop all firms with missing firm identification numbers, province, industry, age, or employment, and drop those with negative values of age or revenue. The final ASIE sample, consisting of 4,037,866 firm-level observations, is the foundation of our firm-level sample.

Step 2: Attach Patent Information to the ASIE Sample. We merge the patent-level Innography Database to the ASIE sample by using information on institutional applicants of patents. Table A.3 presents the number of patents for all and domestic firms. We restrict attention to domestic firms which are favored when Chinese governments give out subsidies (Haley and Haley, 2013).

Table A.3: Number of Patents in the Sample

	All firms				Domestic firms		
Year	Application	Grant	G/A	Firm No.	Grant	% in all	Firm No.
1998	687	412	59.97%	238	364	88.35%	217
1999	1,144	737	64.42%	316	623	84.53%	281
2000	1,831	1225	66.90%	449	1,068	87.18%	408
2001	3,011	2,169	72.04%	610	1,750	80.68%	549
2002	6,272	4,564	72.77%	963	3,066	67.18%	848
2003	9,916	6,360	64.14%	1,357	4,536	71.32%	1,189
2004	14,087	8,343	59.22%	1,799	5,901	70.73%	1,541
2005	19,751	12,124	61.38%	2,291	8,849	72.99%	1,930
2006	28,801	17,274	59.98%	3,192	12,672	73.36%	2,594
2007	37,106	22,122	59.62%	4,068	15,864	71.71%	3,238
2008	46,201	26,871	58.16%	6,071	19,533	72.69%	4,976
2009	53,886	31,431	58.33%	7,303	22,829	72.63%	5,940
2010	72,905	40,133	55.05%	10,422	29,689	73.98%	8,496
2011	98,555	52,366	53.13%	12,266	39,853	76.10%	10,020
2012	136,026	71,726	52.73%	15,565	55,273	77.06%	12,804
2013	165,874	87,304	52.64%	18,790	69,733	79.87%	15,785

Note: The "Grant" column denotes the number of patents that are applied and eventually granted. "% in all" for domestic firms is the fraction of grant patents by domestic firms in all firms.

Step 3: Attach the Firm Innovation Activities Database to the ASIE Sample. For calibration purposes, we further merge a supplementary Firm Innovation Activity Database, available from 2008 to 2013, to the ASIE sample by using firms' organization codes and Chinese names.

A.2.3 Variable Construction

Types of Patents. In the benchmark, we label a Chinese patent as radical if (1) it has been cited by as least one US patent, and (2) the gap of the application year between the cited and citing patent lies within 5 years. Incremental patents are those that are not radical.

Types of Firms. Firms are classified into two types: high- and low-type firms. High-type firms are those with at least one radical patent from 2008-2013; otherwise, firms are labeled as low-type.

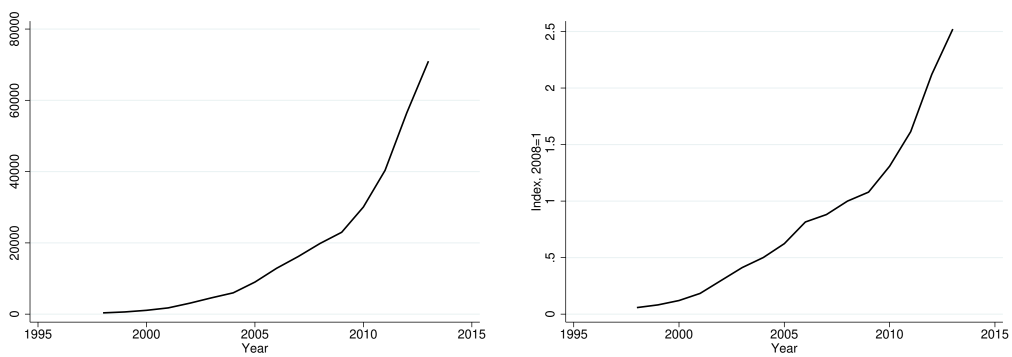
SOE and Foreign Firms. A firm is a state-own enterprise (*resp.* a foreign firm) if either (a) the controlling shareholder is the state (*resp.* foreign or from Hong Kong, Macau, and Taiwan), or (b) the share of state (*resp.* foreign capital and capital from Hong Kong, Macau, and Taiwan) in total capital is greater than 50%.

Skill Composition. We define the employment engaging in scientific activities (*keji huodong renyuan*) as R&D personnel. Among all R&D personnel, we further categorize those with medium or high professional titles (*zhonggaoji zhicheng*) as skilled personnel. Skill intensity is then defined as the ratio of skilled personnel to total R&D personnel.

A.2.4 Additional Tables and Figures

Figure A.4 shows the total number of patents and patents per researcher in our firm-level sample. Same as the aggregate trend shown in Figure 2.1, there appears a speedup in patents per researcher among ASIE firms since the late 2000s. Table A.4 shows the distribution in forward citation codes of Chinese patents. Overall, more than 90% of forward citations for Chinese patents come from other patents applied in China. US patents are the second largest source of citations for Chinese patents, accounting for 7.22% of total forward citations and much larger than the share for other areas.

Figure A.4: Patent Quantity For Industrial Firms



(a) No. of Patent Applications (mn) (b) Patents Per Researcher

Note: This figure shows the number of patents and patents per researcher (index, 2008=1) for ASIE firms. There is no information on the number of researchers for industrial firms before 2011. By assuming that the share of researchers in industrial firms in total researchers is constant, we divide the number of ASIE patents by the number of national researchers, normalize such that the value in 2008 is 1, and present the results in the right panel.

Table A.4: Distribution of Forward Citations for Chinese Patents

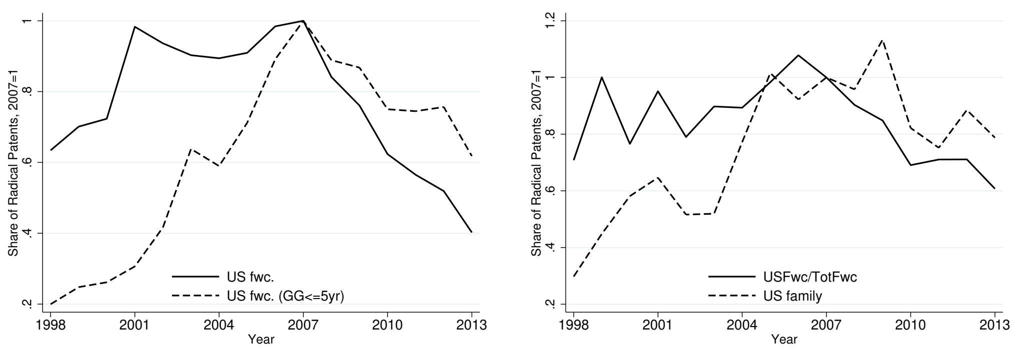
Code	CN	US	TW	JP	EP	KR	GB	DE	AU	FR
Fraction, %	90.73	7.22	0.73	0.44	0.32	0.17	0.07	0.03	0.03	0.03

Data source: Innography and Orbis Intellectual Property

A.2.5 Robustness Check on Patent Quality Decline

Aggregate Trend: Alternative Definitions. Figure A.5 presents the trend of the share of radical patents under four alternative definitions: (a) radical patents are those with at least one forward citation from a US patent; (b) the condition in (a) plus that the gap between grant years of the cited and citing patents are within 5 years; (c), the ratio of No. of US forward citations and No. of all forward citations; (d) radical patents are defined as those with a US patent as its family member. Definition (a) might suffer from truncation bias, while (b)-(d) do not. Though the magnitude varies across different definitions, there displays a rise-then-decline trend in all four series.

Figure A.5: Share of Radical Patents under Alternative Definitions



(a) US forward citation (b) USFwc/TotFwc and US family

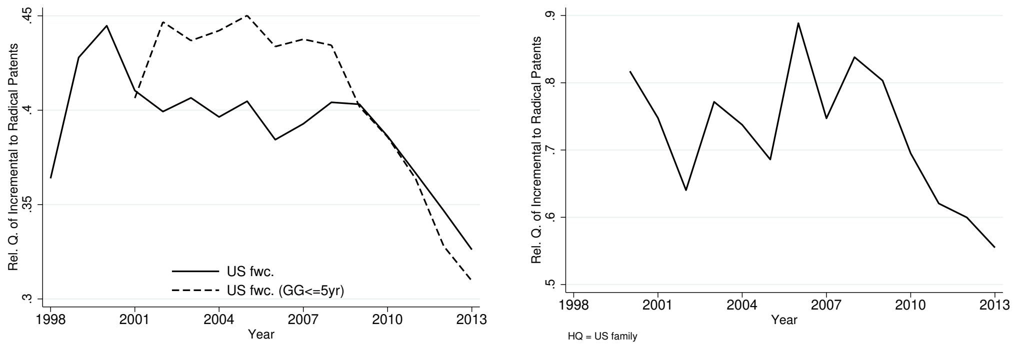
Note: For 'US fwc.', a patent is defined as radical if it is even being cited by a US patent; For 'US fwc. (GG<=5)', a patent is defined as radical if it is ever being cited by a US patent, and the gap in grant year between the cited and citing patent is within 5 years; For 'US family', a patent is defined as radical if it has one US patent as its simple family member; 'USFwc/TotFwc' denotes the fraction of US forward citations in total forward citations.

Figure A.6 presents the trend of the relative quality of incremental to radical patents under definitions (a), (b), and (d),[33] all showing a similar flat-then-decline pattern. In Table A.5, we further restrict to incremental patents and regress the log of forward citations received at the patent level on IPC class-level incremental patent stocks, controlling for IPC class-level radical patent stocks, IPC class, and year fixed effects. The results indicate a

[33]Under definition (c), one cannot properly label whether a patent is radical or incremental.

significantly negative impact of incremental patent stock on incremental patents' quality.

Figure A.6: Relative Quality of Incremental to Radical Patents



(a) US forward citation

(b) US family

Note: See footnote of Figure A.5.

Table A.5: Patent-Level Log Number of Forward Citations Received

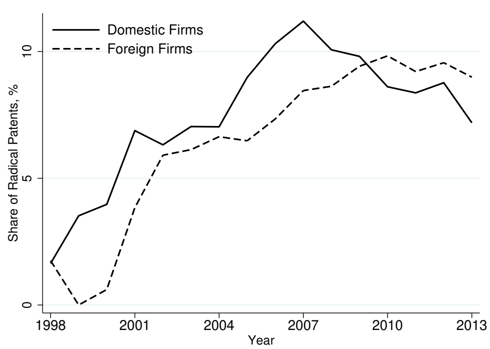
	(1)	(2)	(3)
IPC class-level log incremental patent stock	-0.216***	-0.146***	-0.294***
	(-5.14)	(-4.19)	(-6.64)
IPC class-level log radical patent stock	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Patent Class FE	Yes	Yes	Yes
R^2	0.065	0.069	0.064
Observations	297552	300076	272864

Note: t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table shows results regressing the log of forward citations received at the patent level, of incremental patents, on IPC class-level incremental patent stocks, with IPC class-level radical patent stocks, IPC class, and year fixed effects controlled. Columns [1], [2], and [3] define incremental patents as without US citations within a five-application-year window, without US family applications, and without US citations, respectively. Robust standard errors clustered at the patent class level.

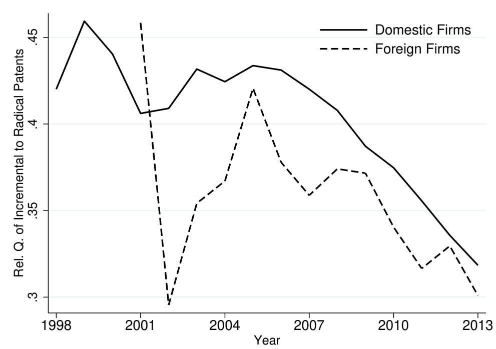
Next, we stick to the baseline definition, i.e., radical patents are cited by at least one US patent, and the gap between the application years of the cited and citing patent is within 5 years, and confirm the robustness of the aggregate trend at more disaggregated levels.

Domestic versus Foreign. Figure A.7 presents the two measures of patent quality for domestic and foreign firms. For radical patent shares (in panel (a)), both firms experience a rising trend in the late 1990s and early-to-middle 2000s. For the post-2008 period, while there is a clear decline for domestic firms, the trend for foreign firms is less visible. That the decline is more significant among domestic firms is consistent with existing research pointing out that China favors indigenous firms when giving out subsidies (Haley and Haley, 2013). As shown in panel (b), the decline in the relative quality of incremental to radical patents is also more significant for domestic firms.

Figure A.7: Evolution of Patent Quality for Domestic and Foreign Firms



(a) Share of Radical Patents (%)



(b) Relative Quality of Incremental to Radical Patents

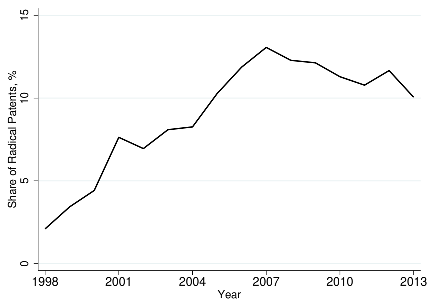
Note: A firm is defined as foreign when the share of registered capital held by foreigners is no less than 50% or when foreigners is the controlling shareholder, and defined as domestic otherwise.

Exporters versus Non-Exporters. A firm is defined as an exporter if it ever exports from 1998 to 2013.[34] Defined this way, exporters account for 65.59% of all patents. Figure A.8 presents the share of radical patents for exporters and non-exporters. The radical patent share for exporters is greater than for non-exporters, which is not surprising since exporters are typically larger in size and more intensive in R&D. The trend for exporters is quite similar to the aggregate one. For non-exporters, the radical patent share does

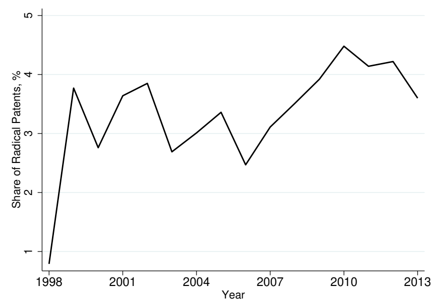
[34]We also tried to define exporters as firms that have an export-revenue ratio exceeding 50% in at least 1 year from 1998 to 2013. The results are qualitatively similar to patterns shown in the text.

not show a clear pre-2008 trend, suggesting a much smaller learning effect for this group. The post-2008 decline for non-exporters is also less significant, though there is a visible decline since the year 2010. Figure A.9 shows the relative quality of incremental to radical patents. Again, the post-2008 decline in relative quality is mainly driven by exporters.

Figure A.8: Share of Radical Patents for Exporters and Non-Exporters



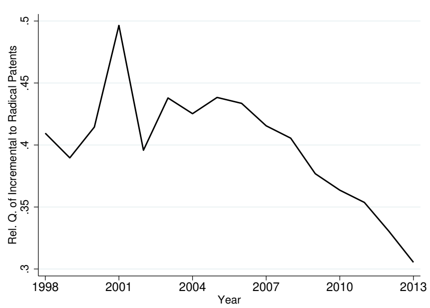
(a) Exporters



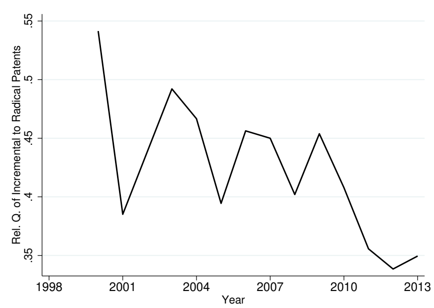
(b) Non-Exporters

Note: A firm is defined as an exporter if it ever exports from 1998 to 2013.

Figure A.9: Relative Quality of Incremental to Radical Patents for Exporters and Non-Exporters



(a) Exporters

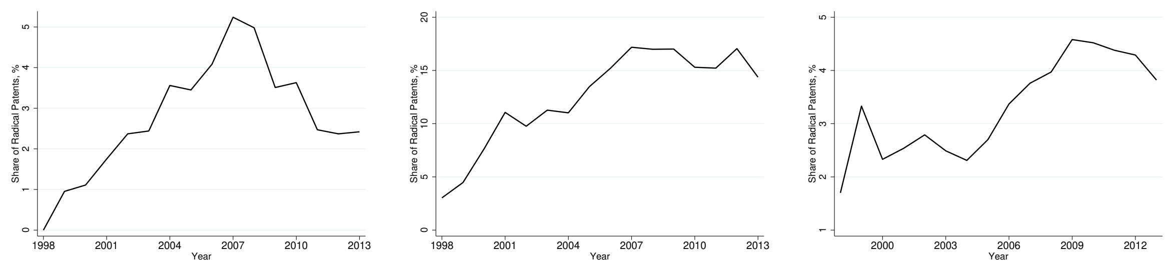


(b) Non-Exporters

Note: see footnote of Figure A.8.

Industry Heterogeneity. Figure A.10 presents the radical patent share for the light, electronics, and heavy industries. There is a clear rise-then-decline pattern for the light and heavy industries. Radical patent share for electronics is roughly stable or shows a mild decline in the post-2008 period. It, however, displays a clear decline compared to the pre-2008 trend. In Figure A.11, we confirm that the relative quality of incremental to radical patents is relatively stable before the mid-to-late 2000s and declines thereafter in all three industries.

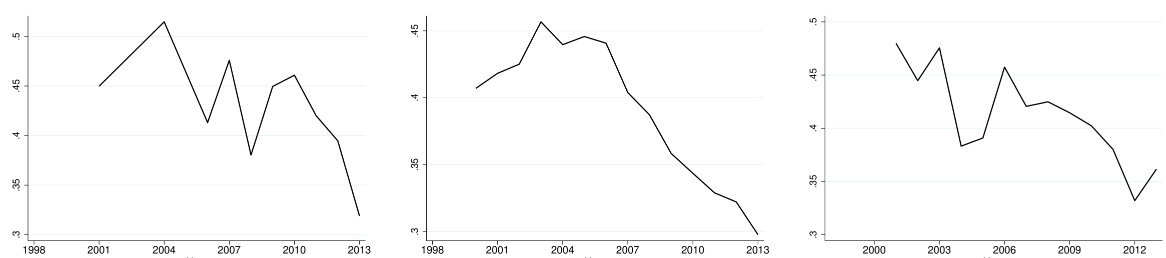
Figure A.10: Share of Radical Patents of Different Industries



(a) Light Industry (b) Electronics Industry (c) Heavy Industry

Note: This figure shows the radical patent share of the light, electronics, and heavy industries.

Figure A.11: Relative Quality of Incremental to Radical Patents of Different Industries



(a) Light Industry (b) Electronics Industry (c) Heavy Industry

Note: This figure shows the relative quality of incremental to radical patents of the light, electronics, and heavy industries.

Different Patent Categories. Table A.6 shows patent numbers and shares for eight 1-digit IPC section symbols, which we refer to as "categories". There are eight categories, with symbols ranging from A to H. Category A refers to "Human Necessities"; Category B refers to "Performing Operation, Transporting"; Category C refers to "Chemistry, Metallurgy"; Category D refers to "Textiles, Paper"; Category E refers to "Fixed Constructions"; Category F refers to "Mechanical Engineering, Lighting, Heating, Weapons, Blasting"; Category G refers to "Physics"; Category H refers to "Electricity". The largest categories are H, C, and B, which account for 30.30%, 17.81%, and 17.20% of total patents created by industrial firms.

Table A.6: No. and Share of Patents for Different Categories, 2000-2013

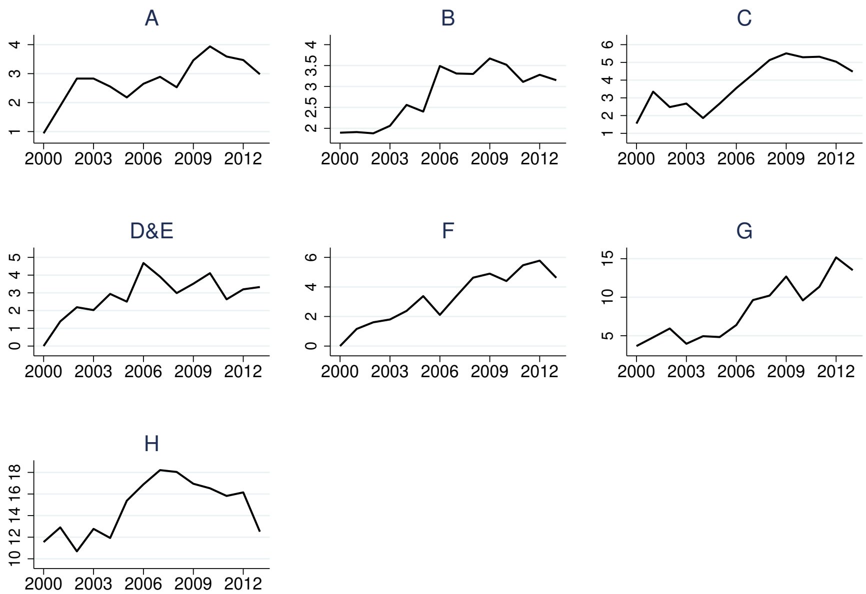
Category	A	B	C	D	E	F	G	H
No.	25,170	50,786	52,582	6,766	7,688	22,234	40,571	89,466
Share	8.52%	17.20%	17.81%	2.29%	2.60%	7.53%	13.74%	30.30%

Note: This table shows the number of patents and the share in total for each patent category (IPC Section Symbols). Category A refers to "Human Necessities"; Category B refers to "Performing Operation, Transporting"; Category C refers to "Chemistry, Metallurgy"; Category D refers to "Textiles, Paper"; Category E refers to "Fixed Constructions"; Category F refers to "Mechanical Engineering, Lighting, Heating, Weapons, Blasting"; Category G refers to "Physics"; Category H refers to "Electricity".

Figure A.12 shows the radical patent share for 7 patent categories (i.e., 1-digit IPC Section Symbols) from 2000-2013.[35] As the number of patents in categories D and E is significantly less than others, containing 6-7 thousand patents in total and less than 500 patents in most of the sample years, we merge those two categories as a single one. For categories A, B, C, D&E, H, or 5 out of the 7 categories, there is a decline in the post-2008 period especially compared to the pre-2008 trend. Together, patents from these five categories account for over 80% of the total patents. As for the relative quality of incremental to radical patents, the flat-then-decline pattern exists for almost all patent classes, as shown in Figure A.13.

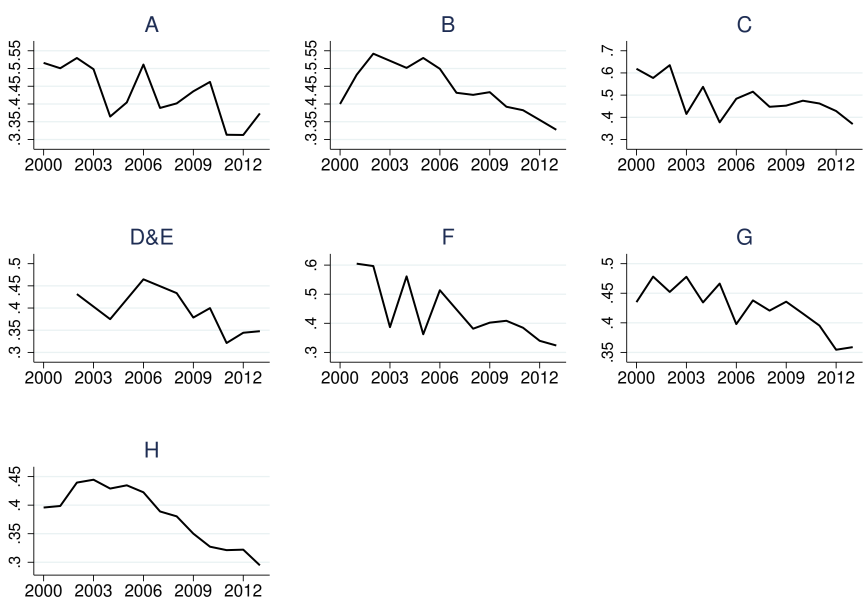
[35]We start from 2000 as there are less than 100 patents for most patent categories in 1998 and 1999. In rare cases that the quality measure for a category in year t is significantly different from year $t-1$ and $t+1$, we interpolate to obtain the value in year t.

Figure A.12: Radical Patent Share for Different Categories



Note: This figure shows the trend of radical patent share for different patent categories (IPC Section Symbols). See note of Table A.6 for detailed meanings of each patent category.

Figure A.13: Relative Quality of Incremental to Radical Patents for Different Categories



Note: This figure shows the trend of the relative quality of incremental to radical patents for different patent categories (IPC Section Symbols). See note of Table A.6 for detailed meanings of each patent category.

Entrants versus Incumbents. A firm is defined as an entrant in year t if its first (eventually-granted) patent is applied in year t, and as an incumbent otherwise. Table A.7 shows the number and share of patents owned by incumbents and entrants from 1998 to 2013. By definition, entrants' patent share equals 100% in 1998. Overall, the share of patents accounted for by entrants does not show a rising trend over the post-2008 period. This result, however, should be interpreted under the caveat that the data we use is for above-scale industrial firms and it does not contain very small firms.

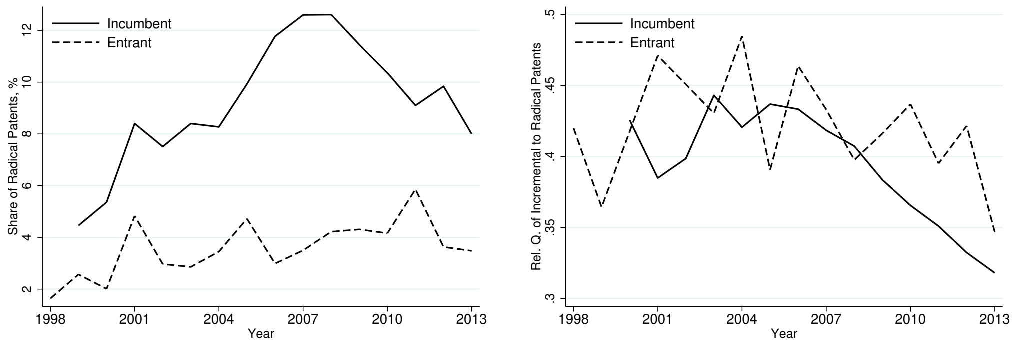
Table A.7: No. and Share of Patents by Entrants

Year	1998	2001	2004	2007	2009	2010	2011	2012	2013
No.	365	746	1,538	2,483	5,286	8,477	9,059	9,712	12,604
Share	100.00%	42.43%	25.69%	15.34%	23.00%	28.20%	22.45%	17.20%	17.75%

Note: A firm is defined as entrant in year t if its first (eventually-granted) patent is applied in year t.

Figure A.14 shows the trend of patent quality for incumbent firms and entrants. As seen in panel (a), the rise-then-decline pattern in radical patent share is predominantly driven by incumbents. For entrants, we actually find a flat or slightly rising trend from 1998-2013. Confining to the later 2008-2013 period, there is no clear decline for entrants. If we take away the jump in 2011, there is a moderate decline for entrants, but with a magnitude much smaller than that of incumbents. In panel (b), the flat-then-decline trend in the relative quality of incremental to radical patents is also predominantly driven by incumbents. No clear trend can be discerned for entrants.

Figure A.14: Evolution of Patent Quality for Incumbent Firms and Entrants



(a) Share of Radical Patents (%)

(b) Relative Quality of Incremental to Radical Patents

Note: A firm is defined as an entrant in year t if its first (eventually-granted) patent is applied in year t.

Regressions With versus Without Firm Fixed Effects. To examine more closely the post-2008 decline in radical patent share, we run a patent-level regression of the radical patent dummy against year (a trend variable) with [column (a)] and without [column (b)] fixed effects. Table A.8 shows the results. The ratio between the coefficient in column (b) and that in column (a) is informative on how much the aggregate decline is explained by the within-firm component. The ratio is 74% for the 2010-2013 period, implying that the post-2009 decline mainly occurs within firms.

Table A.8: Patent-Level Regression with and without Firm Fixed Effects

Period	2010-2013	
	(a)	(b)
Year	-0.0047***	-0.0035***
	(0.0006)	(0.0007)
Firm FE	N	Y
Obs.	197,885	186,101

Note: This table shows regression results in which the dependent variable is the radical patent dummy and the main independent variable is the year trend. Note that firms that appear only once in the sample are automatically dropped in the regression with firm fixed effects.

We also try to use the full sample data from 1998-2013 and regress the radical patent dummy against year dummies with and without firm fixed effects. Then using coefficients for the year dummies, we estimate a pre-2008 trend and compare the relative deviation of the actual coefficients in the post-2008 period from the predicted values extrapolated from the pre-2008 trend. The average deviation from 2011-2013 is 1.07 in the case without firm fixed effects, and 0.71 in the case with firm fixed effects. The latter accounts for 66% of the former, which is consistent with the results in Table A.8.

For decline in the radical patent share that is not "within-firm", one component is the increase in the patent share of low-type firms. By definition, the radical patent share for low-type firms is always zero, so there can be no within-firm decline for these firms. On the other hand, creating more incremental patents by a high-type firm not only reduces the radical patent share within the firm, but also increases its patent share among all firms, therefore simultaneously leading to a within-firm and cross-firm effect, a property featured in the data and captured by our model.

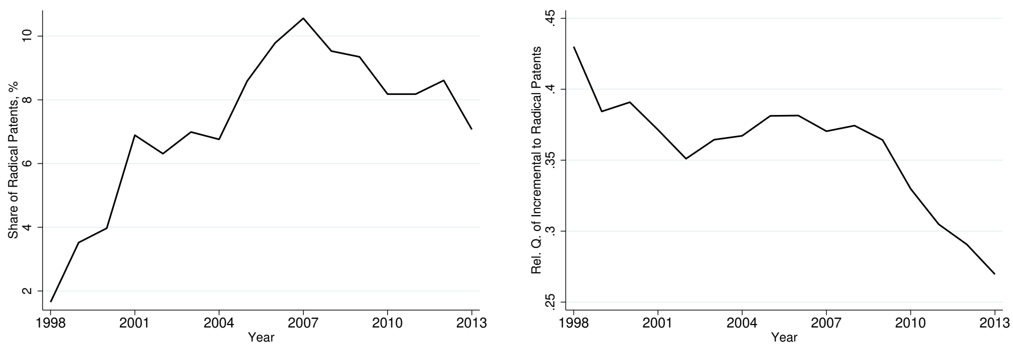
Internal versus External Patents. All (eventually granted) patents are used in the baseline. We follow Akcigit and Kerr (2018) to divide all patents into internal versus external ones based on backward citations. Specifically, we define the self-citation rate as the fraction of *self-citation* in its total backward citations, with *self-citation* of a patent applied by firm i in year t containing all backward citations that are applied between 1998 and year t by firm i. Patents with a self-citation rate greater than 50% are defined as internal patents. External patents are those that are not internal. Patents without backward citations are regarded as external patents. As shown in Table A.9, the vast majority of Chinese patents are external under the definition above. It is not surprising that both the radical patent share and the relative quality of incremental to radical patents among external patents are quite similar to that for all patents, as seen in Figure A.15.

Table A.9: No. and Share of External Patents for Domestic Industrial Firms

Period	No. of external patents	No. of patents	Share of external patents
1998-2013	279,800	296,254	94.45%
2011-2013	159,490	167,823	95.03%

Note: External patents are defined as those with a self-citation rate no greater than 50%.

Figure A.15: Evolution of Patent Quality among External Patents



(a) Share of Radical Patents (%)

(b) Relative Quality of Incremental to Radical Patents

Note: External patents are defined as those with a self-citation rate no greater than 50%.

A.2.6 Skill Intensity for High- and Low-Type Firms

Table A.10 shows the average skill intensity for high- and low-type firms. It is clear that high-type firms are more skill-intensive.

Table A.10: Skill Intensity for Firms of Different Types

	2011	2012	2013	Average
Skill intensity of high-type firms	34.91%	34.53%	33.02%	34.12%
Skill intensity of low-type firms	26.83%	25.19%	24.69%	25.42%

Note: Skill intensity is defined as the fraction of R&D personnel with a medium or senior professional title. To clear the effects of age and size, we first run a firm-level regression of skill intensity against age and log(employment) and then use the residual to obtain the numbers above.

B Appendix: Model Derivations and Proofs

B.1 Education

With the presence of education subsidy b_e and skilled labor subsidy b_h, young people choose to invest in education and become skilled if and only if

$$\frac{e^{-(r-g+d)}}{r-g+d}\frac{w^h}{1-b_h} - (1-b_e)\frac{1-e^{-(r-g+d)}}{r-g+d}\frac{1}{\theta\xi}\frac{w^h}{1-b_h} \geq \frac{w^\ell}{r-g+d},$$

that is, obtaining education if and only if her type is above the threshold

$$\theta^* \equiv \max\left\{\frac{1-b_e}{\xi}\left[1-e^{-(r-g+d)}\right]\left(e^{-(r-g+d)}-(1-b_h)\frac{w^\ell}{w^h}\right)^{-1}, 1\right\}.$$

We can further derive the mass of the four types of people in the economy: students, skilled workers employed in education, skilled workers employed in the R&D sector, and unskilled workers

$$h^{\text{student}} = \theta^{*-2}\left(1-e^{-d}\right)L;$$

$$h^{\text{teacher}} = \frac{2\theta^{*-3}}{3\xi}\left(1-e^{-d}\right)L;$$

$$h^{\text{R\&D}} = \theta^{*-2}e^{-d}L - h^{\text{teacher}};$$

$$\ell^{\text{supply}} = \left(1-\theta^{*-2}\right)L.$$

B.2 Step-Size and Firm Size Distribution

Start with the step-size distribution of incremental innovations. Denote D_τ the fraction of product lines of distance τ, with $\tau = 1$ representing a product line where the latest innovation is radical. Under an invariant distribution,

STATE:	INFLOW		OUTFLOW
$\tau = 1$:	$(1 - D_1)\delta_d$	=	$D_1 \delta_m$
$\tau \geq 2$:	$D_{\tau-1}\delta_m$	=	$D_\tau(\delta_d + \delta_m)$

where δ_d and δ_m are aggregate creative destruction from radical and incremental innovations, respectively. Denote δ the aggregate creative destruction rate, that is, $\delta \equiv \delta_d + \delta_m$. Under the invariant distribution, inflow equals outflow for each $\tau \geq 1$. It follows that,

$$D_\tau = \frac{\delta_d}{\delta}\left(\frac{\delta_m}{\delta}\right)^{\tau-1}, \ \tau = 1, 2, ...$$

From this distribution, we can calculate the expected step-size of an incremental innovation as

$$\bar{\eta} = \sum_{\tau=1}^{\infty} D_\tau \eta \alpha^{\tau-1} = \eta \Big/ \left(\alpha + \frac{1-\alpha}{\delta_d/\delta}\right).$$

For the firm size distribution, denote $p_{\mathrm{H}} = p^*$, $p_{\mathrm{L}} = 1 - p^*$, $x_{\mathrm{H}} = x_{\mathrm{H}d} + x_{\mathrm{H}m}$ and $x_{\mathrm{L}} = x_{\mathrm{L}m}$. Then for firms of type j, stationarity implies that

STATE:	INFLOW		OUTFLOW
$n = 0$:	$\mu_{j,1} \times \delta$	=	$p_j \times x_{\mathrm{E}}$
$n = 1$:	$p_j \times x_{\mathrm{E}} + \mu_{j,2} \times 2\delta$	=	$\mu_{j,1} \times (x_j + \delta)$
$n \geq 2$:	$\mu_{j,n-1} \times (n-1)x_j + \mu_{j,n+1} \times (n+1)\delta$	=	$\mu_{j,n} \times n\,(x_j + \delta)$

For $n = 0$, the inflow occurs when firms with only 1 product line are destroyed, and the outflow is the successful innovations by entrants. For $n = 1$, the inflow contains firms originally with 2 product lines losing 1 line and the entrants who successfully add 1 line; while the outflow consists of 1-line firms that innovate and obtain additional lines or lose existing lines due to creative destruction. A similar interpretation applies for $n \geq 2$. From these expressions, we have

$$\mu_{j,n} = \frac{p_j x_{\mathrm{E}}}{\delta} \left(\frac{x_j}{\delta} \right)^{n-1} \frac{1}{n},$$

and

$$\sum_{n=1}^{\infty} \mu_{j,n} \times n = \frac{p_j x_{\mathrm{E}}}{\delta - x_j}.$$

B.3 Proof of Proposition 1

For a more general theoretical property, let's assume that there are $J \geq 2$ many types of firms in the economy. And define $\mathrm{Line}_j \equiv \sum_n \mu_{j,n} \times n$, that is, the total number of product lines owned by type j firms.

As shown in B.2, stationarity requires that $\forall j \in \{1, 2, ..., J\}$,

$$\mathrm{Line}_j = \frac{p_j x_{\mathrm{E}}}{\delta - x_j}.$$

We plug in the definition of δ and get

$$\mathrm{Line}_j = \frac{p_j x_{\mathrm{E}}}{\sum_{j'} \mathrm{Line}_{j'} x_{j'} + x_{\mathrm{E}} - x_j} \tag{eqn-[j]}$$

with the requirement of

$$\sum_j \text{Line}_j = 1. \tag{eqn-[x]}$$

This is a system of J unknowns $\{\text{Line}_j\}_{j=1}^J$, and $J+1$ equations.

Seemingly, we need an extra free variable such that it is a system of $J+1$ unknowns and $J+1$ equations. However, we are going to prove that, for any given combinations of $x_{\text{E}} > 0$, $\{p_j\}_{j=1}^J \in (0,1)$ and $\{x_j\}_{j=1}^J > 0$, there always exists a $\{\text{Line}_j\}_{j=1}^J$ such that the above $J+1$ equations hold. The remaining is to show that, when eqn-[1] to eqn-[J-1] hold and eqn-[x] is satisfied, the last equation, eqn-[J], shall hold automatically.

Eqn-[j] indicates that

$$(p_j - \text{Line}_j)x_{\text{E}} = \text{Line}_j \left(\sum_{j'} \text{Line}_{j'} x_{j'} - x_j \right),$$

sum them up from j = 1 to $J-1$. Then use the fact $p_J = 1 - \sum_{j=1}^{J-1} p_j$, together with eqn-[x], $\text{Line}_J = 1 - \sum_{j=1}^{J-1} \text{Line}_j$, we have

$$(\text{Line}_J - p_J)x_{\text{E}} = (1 - \text{Line}_J) \left(\sum_{j'} \text{Line}_{j'} x_{j'} \right) - \sum_{j=1}^{J-1} \text{Line}_j x_j.$$

For the R.H.S., we rearrange terms and get

$$(\text{Line}_J - p_J)x_{\text{E}} = (1 - \text{Line}_J)\text{Line}_J x_J - \text{Line}_J \sum_{j=1}^{J-1} \text{Line}_j x_j,$$

which is exactly the same as eqn-[J]

$$(p_J - \text{Line}_J)x_{\text{E}} = \text{Line}_J \left(\sum_{j=1}^{J} \text{Line}_j x_j - x_J \right).$$

B.4 Value Functions and Proof of Proposition 2

As discussed in the main text, we focus on high-type firms which face a trade-off between radical and incremental innovations. Again, for expositional convenience, we drop the firm type subscript j. Guess that the value function takes the following form

$$V(Q,\bar{q}) = \sum_{\omega} Aq_{\omega} + nB\bar{q}.$$

Substituting this conjectured form into the Bellman equation, we have

$$\begin{aligned} r\left(\sum_{\omega} Aq_{\omega} + nB\bar{q}\right) - gnB\bar{q} = \max_{x_d,x_m} \; & \sum_{\omega} \left[\pi q_{\omega} - \delta(Aq_{\omega} + B\bar{q})\right] + nx_d\left[A(1+\lambda) + B\right]\bar{q} \\ & + nx_m\left[A(1+\bar{\eta}) + B\right]\bar{q} - nR(x_d, x_m; z_d, z_m) + nb_n\bar{q}. \end{aligned}$$

It follows that coefficients A and B satisfy the following conditions

$$\begin{aligned} A &= \frac{\pi}{r+\delta}; \\ (r-g+\delta)B &= \max_{x_d,x_m} \; x_d\left[A(1+\lambda)+B\right] + x_m\left[A(1+\bar{\eta})+B\right] - \hat{R}(x_d,x_m) + b_n, \end{aligned}$$

where $\hat{R} \equiv R/\bar{q}$ is the detrended R&D cost per line. One can see that B is increasing in b_n. With the value function's form, equation (10) in Proposition 2 follows immediately from the first-order conditions with respect to x_d and x_m.

If the quantity-based subsidy is posted on the number of new patents instead of the stock, i.e., $nx \times b_x\bar{q}$, the value function and the relation between subsidy and B remain unaltered. The only difference is that "innovation return" now becomes

$$\frac{A(1+\lambda)+B+b_x}{A(1+\bar{\eta})+B+b_x}.$$

C Appendix: Calibration

C.1 External Parameters

Table C.1 summarizes the values of all externally calibrated parameters and their sources.

Table C.1: Externally Calibrated Parameters

Para	Value	Equation	Meaning	Source
ρ	0.02	(1)	time discount rate	literature
ν	3	(1)	intertemporal elasticity of substitution	literature
ϵ	0.22	(2)	E.o.S. in final good production	profitability
L	1		total population	normalization
ϕ	0.49	(3)	innovation elasticity w.r.t. R&D	external estimation
η	$\alpha\lambda$	(5)	initial step-size of incremental inno.	assumption
d	0.03		death rate of the population	years of working
u	25%		corporate tax rate	documentations
b_r	150%		R&D tax credit multiplier	documentations

C.2 Connect Innovation Step-Size to Patents' Number of Forward Citations

In this section, we follow Akcigit and Kerr (2018) and show the map between innovation step-size in the model and the number of forward citations in the patent data. Assume that radical and incremental innovations ($s = \lambda, \eta\alpha^{\tau-1}$) obtain a citation from each subsequent patent with probability $s\kappa$. A radical innovation starts a new technology cluster, which future citations build on, and renders older technology clusters obsolete. Denote $m(\omega, t)$ the number of citable patents in product line ω, and $M(t) = \int_0^1 m(\omega, t)d\omega$ total citable patents in period t. As $m(\omega, t)$ satisfies

$$m(\omega, t + \Delta t) = [m(t) + 1]\delta_m \Delta t + 1 * \delta_d \Delta t + (1 - \delta_m \Delta t - \delta_d \Delta_t)m(t),$$

The law of motion for $M(t)$ is

$$M(\omega, t+\Delta t) = [M(t)+1]\delta_m \Delta t + 1 * \delta_d \Delta t + (1 - \delta_m \Delta t - \delta_d \Delta t) M(t).$$

Imposing $M(\omega, t+\Delta t) = M(\omega, t)$ in steady state, we have $M = \frac{\delta}{\delta_d}$. Denote $\Phi_{\tau,n}$ the fraction of incremental patents with step-size $s_\tau = \eta \alpha^{\tau-1}$ and n citations, among all patents. The inflow and outflow are

STATE:	INFLOW	OUTFLOW
$n = 0$:	$G_{\tau-1}\delta_m =$	$M\Phi_{\tau,0}\delta_d + M\Phi_{\tau,0}\eta\alpha^{\tau-1}\kappa\delta_m$
$n = 1$:	$M\Phi_{\tau,n-1}\eta\alpha^{\tau-1}\kappa\delta_m =$	$M\Phi_{\tau,n}\delta_d + M\Phi_{\tau,n}\eta\alpha^{\tau-1}\kappa\delta_m$

It follows that $\Phi_{\tau,0} = \frac{G_{\tau-1}\delta_m}{M\delta_d + M s_\tau \kappa \delta_m}$, and $\Phi_{\tau,n} = \Phi_{\tau,0}(\frac{s_\tau \kappa \delta_m}{\delta_d + s_\tau \kappa \delta_m})^n$. Note that at any point in time, the **number** of $\tau - th$ incremental patents is $G_\tau = M \frac{\delta_d}{\delta} (\frac{\delta_m}{\delta})^\tau$. Then at this point of time, the **fraction** of incremental patents that have citation n, among all incremental patents, is

$$\tilde{f}(n; \alpha, \eta\kappa) = \frac{\sum_{\tau=1}^{\infty} M\Phi_{\tau,n}}{\sum_{\tau=1}^{\infty} G_\tau} = \sum_{\tau=1}^{\infty} \frac{\delta_d}{\delta} \left(\frac{\delta_m}{\delta}\right)^{\tau-1} \frac{\delta}{\delta_d + s_\tau \kappa \delta_m} \left(\frac{s_\tau \kappa \delta_m}{\delta_d + s_\tau \kappa \delta_m}\right)^n$$

Matching this distribution with that from data produces an estimate for α and $\eta\kappa$.

The average citation for $\tau - th$ incremental patents is $\sum_{n=0}^{\infty} \Phi_{\tau,n} n / \sum_{n=0}^{\infty} \Phi_{\tau,n} = s_\tau \kappa \delta_m / \delta$. It follows that the average citation for all incremental patents is $\bar{\eta} \kappa \delta_m / \delta$.

Similarly, denote $\Phi_{\lambda,n}$ the fraction of radical innovations with n citations. The associated inflow and outflow are

STATE:	INFLOW		OUTFLOW
$n=0$:	δ_d	=	$M\Phi_{\lambda,0}\delta_m + M\Phi_{\lambda,0}\lambda\kappa\delta_m$
$n=1$:	$M\Phi_{\lambda,n-1}\left[\lambda\kappa\delta_m + \psi\lambda\kappa\delta_d\right]$	=	$M\Phi_{\lambda,n}\delta + M\Phi_{\lambda,n}\lambda\kappa\delta_m$

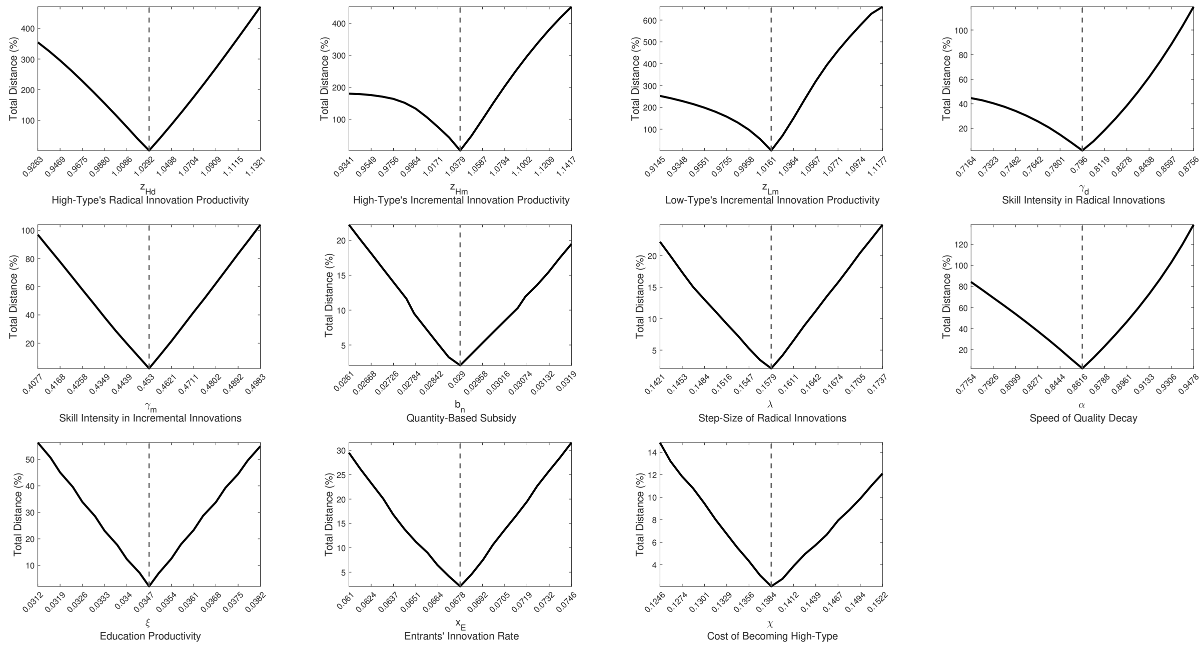
We then have that $\Phi_{\lambda,0} = \frac{\delta_d}{M\delta_m + M\lambda\kappa\delta_m}$ and $\Phi_{\lambda,n} = \left[\frac{\lambda\kappa\delta_m}{\delta_m+\lambda\kappa\delta_m}\right]^n \Phi_{\lambda,0}$. The average citation for radical innovations are $\sum_{n=0}^{\infty} \Phi_{\lambda,n} n / \sum_{n=0}^{\infty} \Phi_{\lambda,n} = \frac{\lambda\kappa\delta_m}{\delta}$. Therefore, we can use the ratio between average forward citations received by incremental and radical patents to infer the $\bar{\eta}/\lambda$ ratio.

C.3 Identification

To formally illustrate the identification of internally calibrated parameters, we conduct two exercises. First, we show how the total sum of distance changes as we move one parameter away from its benchmark value while keeping the others unchanged. Figure C.1 summarizes the results. One can check that the total distance is well V-shaped with respect to all parameters, with its minimum achieved at the benchmark value (dash line), which implies that the identification is clear.

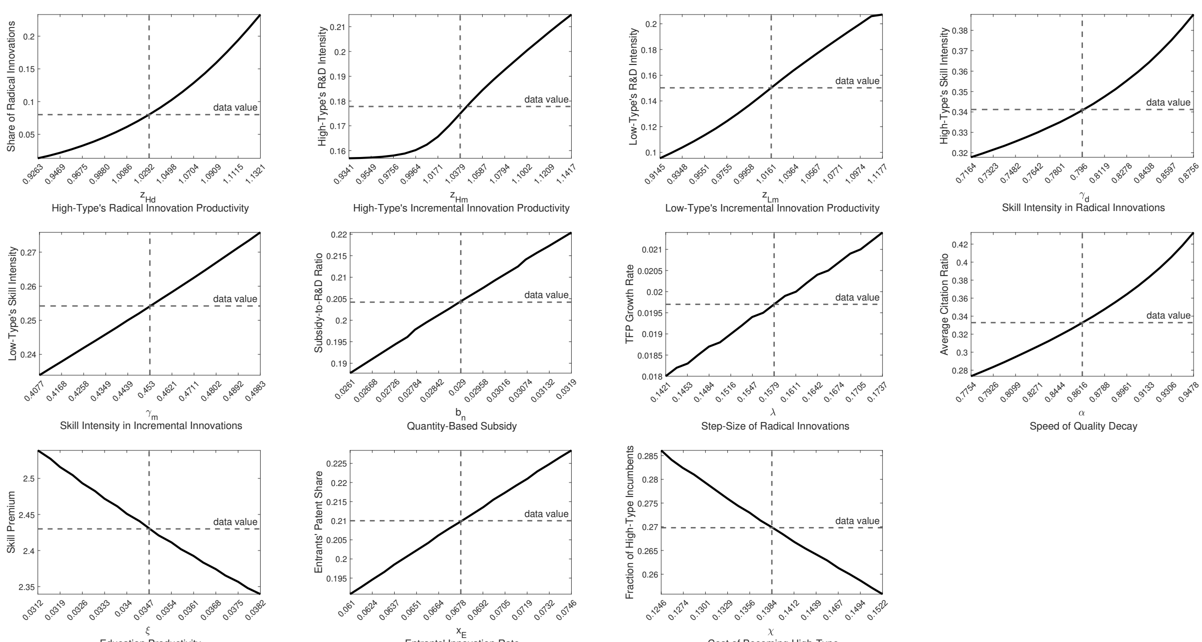
Second, in the main text, we picked a most informative moment for each of the parameters based on model implications, despite the fact that all 11 parameters are jointly identified. To support this argument, we check each of the 11 model-generated moments as a function of the corresponding parameter. Figure C.2 shows the results, the identification is clear as well.

Figure C.1: Total Distance w.r.t. Each Parameter



Note: This figure shows how the total sum of distance (Y-axis) changes as we move each of the 11 parameters (X-axis) away from its benchmark value, up and down by 10%, while keeping the others unchanged.

Figure C.2: Informative Moment w.r.t. Each Parameter



Note: This figure checks the sensitivity of each of the 11 model-generated moments (Y-axis) as a function of the corresponding parameter (X-axis). Again, each parameter is moved up and down by 10% from its benchmark value, while keeping the others unchanged.

C.4 More on Model Fit: Patent Stock, Relative Firms Size and the Quantity-Quality Trade-off

Active Patent Stock. We construct the stock of active patents using patents' forward citation information. More specifically, we define the lifespan of a patent as the period from its application year to the last year it receives a forward citation. The idea here is essentially to regard an old patent as "inactive" or "dead" when it no longer contributes to society's knowledge creation.

For an eventually granted patent that was applied in year t_0, if the application year associated with its latest forward citation is t, then this patent is treated as "active" between year t_0 and t. If a patent does not receive any forward citation, we assume that it is active only in its application year and becomes inactive in the subsequent periods. For any given year, we then construct the patent-level creative destruction rate as the ratio of newly granted patents and the active patent stock. As the active patent stock grows rapidly, we tried the active patent stock in the current year, the active patent stock in the previous year, and the two-year moving averages of the active patent stock in the two consecutive years, which gives us a range of values. The average creative destruction rate ranges from 30% to 34% in the 2011-2013 period. Since we look at patent-level creative destruction and China experienced a patent surge in that period, we consider the relatively high rate reasonable.

Relative Firm Size. The model predicts that firms with a higher innovation intensity have a larger expected size. Under the calibrated parameter values, the size ratio between average high- and low-type firms, measured by employment, revenue, or profit, is 1.332. Table C.2 shows the relative size ratio from 2011-2013 firm-level data. Our calibration captures the size difference well.

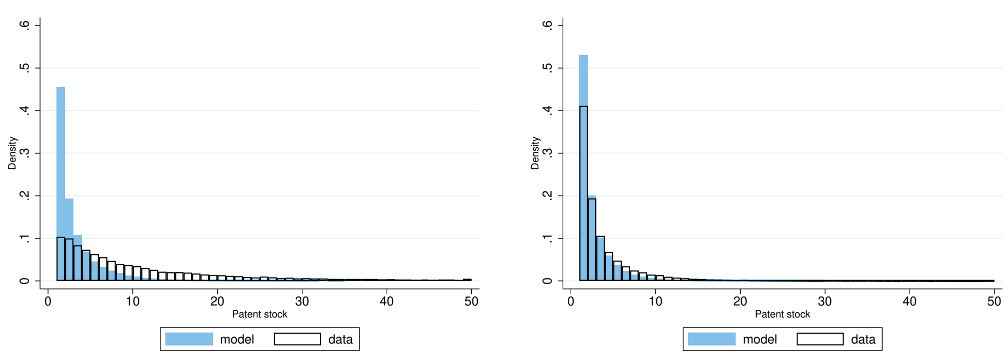
Table C.2: Size Ratio between High- and Low-Type Firms in the Data and the Model

	Employment	Revenue	Profit
Data	1.139	1.249	1.420
Model	1.332	1.332	1.332

Note: This table reports the relative ratio for variables of interest, between average high- and low-type firms in the 2011-2013 period, and we trim the bottom and the top 5 percent of the sample.

We then check the model's performance on the number of patents distributed among innovating firms. In the model, the number of patents corresponds to the number of product lines, n. Similar to the measurement of patent-level creative destruction, we calculate the patent stock of an individual firm in 2011-2013 by summing all active patents. Figure C.3 shows the distribution of active patent stock among high- and low-type firms. We underestimate the number of patents for high-type firms, but overall, the model matches the data pattern well.

Figure C.3: Distribution of Patent Number among High- and Low-Type Firms



(a) High-Type Firms

(b) Low-Type Firms

Note: This figure shows the distribution of patent numbers among high- (panel (a)) and low-type (panel (b)) firms. Patent stock is calculated as the sum of all active patents within the 2011-2013 period, and the distribution of patent stock is then estimated for the two sub-groups.

Magnitude of Quantity-Quality Trade-off. We provide evidence that the model-implied magnitude of quantity-quality trade-off under the calibrated parameters is in line with data, by exploiting firm-level variations from the Innocom Program. China initiated many policies that aim to promote firm innovations around the mid-2000s. A critical subsidy program China has initiated to promote firm innovations around the mid-2000s is the recognition of High-Tech Enterprises (HTEs) under the InnoCom Program.[36] Certified HTEs enjoy corporate tax cuts and various types of research and development subsidies such as research grants and patent subsidies, which presumably affect their choices over radical and incremental patents.

We employ a Difference-in-Difference (DID) methodology to examine the impact of HTE recognition on high-type firms' innovation choices.[37] Our approach involved defining a post dummy variable equal to 1 for an HTE firm if the observation year was on or after the year the firm obtained the HTE title for the first time, and 0 otherwise. In contrast, for a Non-HTE firm, the post dummy is always 0. Furthermore, we only considered HTEs with at least one prior year and one post year (including the recognition year) to enable meaningful before-after comparisons. We then regress (log) the share of radical patents on the HTE dummy, the interaction term between the HTE dummy and post dummy, which is the key variable of our interest, controlling for (log) employment, (log) revenue, (log) assets as well as year, location (province), industry, ownership types and established year fixed effects. Robust standard errors are clustered at the firm level.

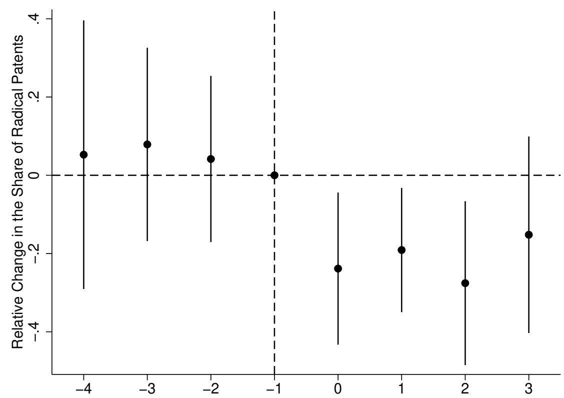
Our DID regression analysis reveals a 24.10% decline in its share of radical patents after a high-type firm receives HTE recognition. We further conduct an event study using

[36] Among the qualifications to become an HTE, the most important criteria are: (1) firms own patents on their core technology and use such core technology on their main production lines where patents can be invented, transferred, purchased or via M&A, (2) R&D related personnel is no less than 10% of the employers, and (3) depending on the level of total sales, R&D expenses must reach a certain amount.

[37] By definition, low-type firms are those who create only incremental patents, so their radical patent share is unaffected.

year zero as the treatment year to confirm the parallel trends in the pre-treatment period. Figure C.4 displays the results, indicating no apparent trends during the pre-treatment periods. However, during the post-treatment periods, we observed a clear decline in the firm's radical patent share, with most of the point estimates being significant.[38]

Figure C.4: Event Study



Correspondingly, we extend the model to include four types of firms: high-type & HTE, high-type & non-HTE, low-type & HTE, low-type & non-HTE. On top of the general quantity-based subsidies eligible for all, HTE firms are rewarded more: $b_n^{\text{HTE}} > b_n^{\text{non-HTE}}$. Consistent with the data, we give each type a 55% chance of HTE recognition. We then quantify the extended model's two extra parameters, $b_n^{\text{HTE}} = 0.034$ and $b_n^{\text{non-HTE}} = 0.019$, according to the observed subsidy differences between HTE and non-HTEs, while keeping the remaining parameters at their benchmark values. The extended model generates a 26.55% decline in its share of radical innovations once a high-type firm receives HTE recognition, quite close to that in the data.

[38]To maintain a sufficiently large sample size for our DID regression and event study, we choose not to apply the 5-year restriction in application years between the cited and citing patents. This should not cause bias as in the regression we include year fixed effects, which resolves the truncation issue. If we had imposed the 5-year restriction, though, one-third of the observations would've been lost, and the DID coefficient became -18.7%, significant at 5% level.

D Appendix: Quantitative Analysis

D.1 Estimating the Magnitude of Patent Quantity Surge and Quality Decline

As shown in Figure A.4, the quantity of patents per researcher increases at a faster rate in the post-2008 period than in the pre-2008 period, while panel (a) in Figure 2.2 shows a clear post-2008 decline in the share of radical patents comparing to the pre-2008 trend. To estimate the magnitude of patent quantity increase above, and of radical share below, their natural trends in the post-2008 period, we first fit the pre-2008 data with a linear trend and then use that trend to extrapolate to obtain the "natural" level for years after 2008 trend. Then by calculating the deviation of the actual level from 2011 to 2013 from the predicted values in relative terms, we obtain the estimation of the magnitude of patent surge above the trend, and of radical share decline below the trend. The estimated quantity increase and radical share decline are 34.57% and 40.89% respectively. For the relative quality of incremental to radical patents (panel (b) in Figure 2.2), the pre-2008 linear trend turns out to be insignificant, so we calculate a 20.27% decline as the relative change of the 2011-2013 average from the 2006-2008 average.

D.2 Key Parameters for the Growth and Welfare Implications

The baseline model finds a negative growth and welfare effect of quantity-based subsidies. The assumption of scarce research time, and parameter values about the degree of quality decay and heterogeneity in skill intensities, determine the strength of quality and quantity channels and the sign of the net impact. We counterfactually shut down each of the three margins to illustrate its importance on the findings presented in Section 4.2. Table D.1 summarizes the results.

Table D.1: Pre vs. Post Changes when Corresponding Margins are Shut Down

Parameter	**Margin**	$\Delta_{\delta - x_{\mathrm{E}}}$	$\Delta_{\delta_d/\delta}$	$\Delta_{\bar{\eta}/\lambda}$	Δ_g	$\Delta_{\textbf{welfare}}$
e	research time	6.36%	-0.24%	-0.06%	0.21 p.p.	0.29%
α	quality decay	9.34%	-41.58%	0.00%	-0.05 p.p.	-1.92%
γ_d, γ_m	skill intensity	8.77%	-12.49%	-8.36%	-0.04 p.p.	-1.79%
baseline results		10.14%	-22.91%	-15.28%	-0.19 p.p.	-3.31%

Note: Δ represents changes in the variable from the counterfactual (without b_n) to the benchmark economy (with b_n) when the corresponding margin is shut down. We present Δ_g in absolute percentage point (p.p.) changes, while the others are presented in relative percentage changes (%).

The first margin regards research time, e. As indicated by the R&D cost function, the scarceness of e helps generate a strong firm-level quantity-quality trade-off. Once we shut it down, i.e., taking e out from the R&D production function, both quality channels are substantially weakened, and the growth and welfare implications are completely reversed (row 1).

The second margin regards innovation quality decay, α. From the growth decomposition in Table 4.4, we already demonstrated that the negative quality-crowding channel is quantitatively dominant. Slightly different from what we did with that decomposition, here we shut down the quality decay margin by setting $\alpha = 1$ and fixing $\eta = 0.053$, i.e., the baseline level of $\bar{\eta}$. As a result, the quality-crowding channel is eliminated, and the growth and welfare implications are largely weakened (row 2).

The last is a pair of parameters regarding skill intensity, γ_d & γ_m. As explained by Proposition 2, the skill intensity difference is through which the general equilibrium skill premium effect works. We shut down this margin by setting $\gamma_d = \gamma_m = 0.453$, that is, the baseline value of γ_m.[39] To make a reasonable comparison, we also adjust $z_{\mathrm{H}d} = 0.970$

[39]Results remain similar if we set γ_d and γ_m to the baseline value of $\gamma_d = 0.796$, so we don't present them here.

to restore the baseline level of radical innovation share δ_d/δ. Consequently, both quality channels are largely weakened, and so are the growth and welfare implications (row 3).

Other parameters, or margins, may also affect the growth and welfare implications to some extent. For example, the entrant's overhead investment, or the R&D tax credit multiplier. However, their effects are secondary compared to the three key parameters mentioned above.

D.3 Extension: Decreasing Return to Scale

In the baseline model, innovation cost scales up linearly with firm size, so that size of a firm does not impact its innovation intensities. While this simplification delivers a clean characterization of the firm-level quantity-quality trade-off and facilitates aggregation, it might lead to bias in policy evaluations. To alleviate this concern, we extend the baseline model to decreasing return to scale (D.R.S.) and re-evaluate the effects of quantity-based innovation subsidies.

Setting. We assume a more general innovation production function with size-dependent R&D productivity

$$X_i = z_i(n)\, n^{1-\phi}\left(e_i\, h_i^{\gamma_i}\ell_i^{1-\gamma_i}\right)^{\phi},$$

where $z_i(n) = z_i n^{-\psi_i}$, for $i = d, m$. Parameters ψ_d, $\psi_m > 0$ govern the speed of productivity decay w.r.t. firm size.[40] Similarly, we can derive the function of R&D cost per line

$$R(x_d, x_m; n) = \left[\Theta_d(n)^{\frac{1}{2}} + \Theta_m(n)^{\frac{1}{2}}\right]^2,$$

but Θ_d and Θ_m are now size-dependent

[40] This specification is effectively assuming $X_i = z_i\, n^{1-\phi-\psi_i}\left(e_i\, h_i^{\gamma_i}\ell_i^{1-\gamma_i}\right)^{\phi}$, for $i = d, m$.

$$\Theta_i(n) = \Delta_i \left(w^h\right)^{\gamma_i} \left(w^\ell\right)^{1-\gamma_i} \left(\frac{x_i}{z_i(n)}\right)^{\frac{1}{\phi}}, \text{ for } i = d, m.$$

Now that the R&D cost per line is size-dependent, the value function takes the form

$$V(Q, \bar{q}) = \sum_{\omega} A q_\omega + B_n \bar{q}.$$

To make a comparison, our benchmark case is where $B_n = n \times B$.

Again, it is easy to verify that

$$A = \frac{\pi}{r + \delta},$$

while the sequence B_n solves

$$\begin{aligned}(r-g)\frac{B_n}{n} = \ & \delta(B_{n-1} - B_n) + b_n \\ & + \max_{x_d, x_m} \ x_d \left[A(1+\lambda) + B_{n+1} - B_n\right] \\ & + x_m \left[A(1+\bar{\eta}) + B_{n+1} - B_n\right] - \hat{R}(x_d, x_m, n),\end{aligned}$$

where $\hat{R} \equiv R/\bar{q}$ is the detrended R&D cost per line. Lastly, a similar but generalized version of Proposition 2 can be derived, that is

$$\frac{x_d(n)}{x_m(n)} \propto \underbrace{\frac{A(1+\lambda) + B_{n+1} - B_n}{A(1+\bar{\eta}) + B_{n+1} - B_n}}_{\text{innovation return}} \times \underbrace{\left(\frac{w^h}{w^\ell}\right)^{-(\gamma_d - \gamma_m)}}_{\text{input structure}} \times \underbrace{\frac{z_d(n)}{z_m(n)}}_{\text{R\&D productivity}}.$$

Here we have an extra "R&D productivity" term, as the relative productivity ratio is no longer a constant. In conclusion, the change does not alter the quantity-quality trade-off facing innovating firms, except that the magnitude now depends on firm size.

Calibration. We assume identical ψ_m for both high- and low-type firms to reduce parameters. To discipline the values of ψ_d and ψ_m, we first estimate two regressions regarding the elasticity of innovation quantity and quality w.r.t. firm size

$$\frac{New\ Patent_{f,t}}{emp_{f,t}} = \beta_0 - \underbrace{0.0411^{***}}_{(s.e.\ 0.0011)} \times \ln(emp_{f,t}) + FE^1_{f,t} + \varepsilon_{f,t};$$

$$Radical\ Patent\ Share_{f,t} = \beta_0 - \underbrace{0.0293^{***}}_{(s.e.\ 0.0024)} \times \ln(emp_{f,t}) + FE^2_{f,t} + \varepsilon_{f,t},$$

where $FE^2_{f,t}$ contains year and industry fixed effects, and $FE^1_{f,t}$ further includes a dummy for high-type firms. We then calibrate the values of ψ_d and ψ_m to match the two elasticity coefficients in the model and the data.

Results. The calibrated $\psi_d = 0.061$ and $\psi_m = 0.055$, which suggest a rather mild D.R.S. among Chinese innovating firms, comparing to the values, $\psi_d = \psi_m = 0.105$, reported in Akcigit and Kerr (2018) regarding US firms. For better comparison with the baseline case, we also apply a common factor $\psi_z = 1.066$ to scale up values of R&D productivity, to restore the baseline level of aggregate creative destruction rate. All the rest of the parameters are kept at their baseline values. Table D.2 & D.3 replicate the growth and welfare implications of quantity-based subsidies, i.e., Table 4.3 & 4.4 in the main text. The results are quite close to the baseline case.

Table D.2: Impact of Quantity-based Subsidies on Innovation Quantity and Quality under D.R.S.

Variable	Meaning	Model	C.F.	Δ_{Model}	Δ_{Data}	$\frac{\Delta_{\text{Model}}}{\Delta_{\text{Data}}}$
$\delta - x_E$	incumbent innovation	25.43%	23.16%	9.80%	34.57%	28.35%
δ_d/δ	radical share	7.37%	9.51%	-22.50%	-40.89%	55.03%
$\bar{\eta}/\lambda$	step-size ratio	31.45%	37.20%	-15.46%	-20.27%	76.27%

Note: Δ_{Model} represents changes from the counterfactual to the model benchmark, Δ_{Data} is changes between the pre- and post-2008 period, both columns are presented in relative terms. Step-size ratio denotes the relative step-size of incremental to radical innovations.

Table D.3: Growth Decomposition under D.R.S.

Δ_{Growth}	(i) quantity	(ii) quality-composition	(iii) quality-crowding	
-0.18	0.15	-0.06	-0.24	(p.p.)
	-83.33%	33.33%	133.33%	

Note: For each of the channels, we add the corresponding change in (i) δ; (ii) δ_d/δ; (iii) $\bar{\eta}$ to the pre-2008 economy, and see how it affects the aggregate growth rate. The second row shows the contribution of each channel, calculated by dividing the corresponding number by -0.18 p.p.

D.4 Extension: Internal Innovations

In the baseline model, all innovations (patents) trigger creative destruction, i.e., being external. We made this assumption based on the observation that the vast majority of Chinese patents are external ones (Table A.9). In this section, we explore how the model would perform when internal innovations are allowed.

Setting. We add a third *internal* innovation choice for all incumbent firms, high- or low-type. In particular, internal innovations arrive at the following Poisson flow rate

$$X_s = z_s n^{1-\phi} \left(h_s^{\gamma_s} \ell_s^{1-\gamma_s} \right)^{\phi},$$

where the subscript s stands for "self-improvement". The arrival of an internal innovation improves the quality of a product line owned by the firm by a fixed step-size $\lambda_s > 0$. The setup gives us an R&D cost per line similar to that assumed in Akcigit and Kerr (2018)

$$R(x_d, x_m, x_s; z_d, z_m, z_s) = \underbrace{\left[\Theta_d(x_d)^{\frac{1}{2}} + \Theta_m(x_m)^{\frac{1}{2}} \right]^2}_{\text{external innovation cost}} + \underbrace{\Theta_s(x_s)}_{\text{internal innovation cost}},$$

where $\Theta_i(x_i) \equiv \Delta_i \left(w^h\right)^{\gamma_i} \left(w^\ell\right)^{1-\gamma_i} (x_i/z_i)^{\frac{1}{\phi}}$ and $\Delta_i \equiv \gamma_i^{-\gamma_i}(1-\gamma_i)^{\gamma_i-1}$, for $i = d, m, s$.

The value function still takes the form $V(Q, \bar{q}) = \sum_{\omega} A q_\omega + nB\bar{q}$, and the return for internal innovations is $A\lambda_s$. Quantity-based subsidies affect the relative return of internal vs.

external innovations through competing forces, e.g., B and $\bar{\eta}$. The overall effect, however, is an issue we address quantitatively.

Before quantifying the model, we introduce some extra notations regarding internal innovations and aggregation. Denote ι_s the aggregate internal innovation rate, we have

$$\iota_s = \sum_j \sum_n \mu_{j,n} \times n x_{js},$$

where x_{js} denotes the internal innovation intensity of the type $j = \mathrm{H}, \mathrm{L}$ firm.

The economy's aggregate innovation rate is $\iota = \delta_d + \delta_m + \iota_s$, while $\delta = \delta_d + \delta_m$ still denotes the aggregate creative destruction rate. The aggregate growth rate is given by

$$g = \delta_d \lambda + \delta_m \bar{\eta} + \iota_s \lambda_s.$$

We follow the baseline approach to decompose changes in g into three channels

$$\Delta g = \underbrace{\Delta\iota \times \left(\frac{\delta_d}{\iota}\lambda + \frac{\delta_m}{\iota}\bar{\eta} + \frac{\iota_s}{\iota}\lambda_s\right)}_{\text{(i) quantity}} + \underbrace{\iota \times \frac{\delta}{\iota} \times \left[\Delta\frac{\delta_d}{\delta} \times (\lambda - \bar{\eta})\right] + \iota \times \Delta\frac{\delta}{\iota} \times \left(\frac{\delta_d}{\delta}\lambda + \frac{\delta_m}{\delta}\bar{\eta} - \lambda_s\right)}_{\text{(ii) quality-composition}}$$
$$+ \underbrace{\iota \times \frac{\delta}{\iota} \times \left[\left(1 - \frac{\delta_d}{\delta}\right) \times \Delta\bar{\eta}\right]}_{\text{(iii) quality-crowding}}.$$

The *quantity* channel now refers to change in the aggregate innovation rate ι; the *quality-composition* channel summarizes changes in the weights δ_d/δ and δ/ι;[41] the *quality-crowding* channel still refers to the change in the average productivity impact of incremental innovations, $\bar{\eta}$.

[41] The former is the weight of radical innovations within external innovations, while the latter is the weight of external innovations among total innovations. Under our calibration, the *quality-composition* effect is mainly driven by its first component.

Calibration. We assume identical γ_s, z_s, λ_s for both high- and low-type firms to reduce the number of parameters. We further assume $\gamma_s = \gamma_m$, i.e., internal innovations have the same skill intensity as the external incremental ones. That leaves us with two parameters, z_s and λ_s, which we calibrate jointly by matching the share of internal innovations and the average citation ratio between internal and external radical patents. For better comparison with the baseline case, we also apply a common scale factor ψ_z on $z_{\mathrm{H}d}$, $z_{\mathrm{H}m}$ and $z_{\mathrm{L}m}$, to restore the baseline level of aggregate innovation rate. All the rest of the parameters are kept at their baseline values.

Results with CN Internal Patent Share. We first take the model to the Chinese patent data. In 2011-2013, the share of internal patents in China is 4.97%, and the average citation ratio between internal and external radical patents is 0.403. These moments produce an internal innovation productivity of $z_s = 1.294$ and a step-size of $\lambda_s = 6.36\%$. The common scale factor required to restore the baseline aggregate innovation rate is $\psi_z = 0.951$.

Table D.4 reports the impact of subsidies on innovation quantity and quality, with an extra row documenting change in the share of internal innovations, while Table D.5 reports the impact on growth.[42]

The results remain very close to the baseline case, reaffirming our strategy of focusing on external innovations in the Chinese context. Moreover, the extended model is able to replicate the decline in the share of internal patents in China, which drops from 6.45% pre-2008 to 4.97% post-2008, or a relative decline of 22.95%. In our quantified model, the introduction of quantity-based subsidies causes firms to focus more on external innovations, which crowd out internal innovations through general equilibrium pricing effects.

[42] In the Chinese sample, the radical patent share within external patents is rather close to the baseline where we do not distinguish internal versus external patents, hence we stick to its baseline values for the convenience of comparison.

Table D.4: Impact of Quantity-based Subsidies on Innovation Quantity and Quality with CN Internal Patent Share

Variable	Meaning	Model	C.F.	Δ_{Model}	Δ_{Data}	$\frac{\Delta_{\text{Model}}}{\Delta_{\text{Data}}}$
$\iota - x_{\text{E}}$	incumbent innovation	25.55%	23.46%	8.91%	34.57%	25.77%
δ_d/δ	radical within external	8.08%	10.45%	-22.68%	-40.89%	55.47%
$\bar{\eta}/\lambda$	step-size ratio	33.47%	39.42%	-15.09%	-20.27%	74.44%
ι_s/ι	internal share	4.97%	5.79%	-14.16%	-22.95%	61.70%

Note: Δ_{Model} represents changes from the counterfactual to the model benchmark, Δ_{Data} is changes between the pre- and post-2008 period, both columns are presented in relative terms. Step-size ratio denotes the relative step-size of incremental to radical innovations.

Table D.5: Growth Decomposition with CN Internal Patent Share

Δ_{Growth}	(i) quantity	(ii) quality-composition	(iii) quality-crowding	
-0.18	0.15	-0.07	-0.24	(p.p.)
	-83.33%	38.89%	133.33%	

Note: For each of the channels, we add the corresponding change in (i) ι; (ii) δ_d/δ and δ/ι; (iii) $\bar{\eta}$ to the pre-2008 economy, and see how it affects the aggregate growth rate. The second row shows the contribution of each channel, calculated by dividing the corresponding number by -0.18 p.p.

All the growth-decomposition channels are weakened, so as the growth and welfare loss from quantity-based subsidies. The magnitude, though, is rather small, as the share of internal innovations in China is negligible. Many Chinese firms in that period of time had very few patents of their own to cite. Mechanically, this leads to a low internal patent share. We leave a detailed discussion on why the growth-decomposition channels are weakened in the following exercise with the US internal patent share.

Results with US Internal Patent Share. Due to the low internal patent share in China, the quantified model generates very similar results to that in the baseline case. As a comparison, Akcigit and Kerr (2018) reports that the US internal patent share is 21.5%. We further conduct a numerical exercise, targeting the share of internal innovations at the US level. This gives a higher internal innovation productivity of $z_s = 2.506$, and a common scale factor of $\psi_z = 0.783$ to restore the baseline aggregate innovation rate. Tables D.6 and D.7 report the impact of subsidies on innovation quantity and quality, and on growth.

Table D.6: Impact of Quantity-based Subsidies on Innovation Quantity and Quality with US Internal Patent Share

Variable	**Meaning**	**Model**	**C.F.**	Δ_{Model}	Δ_{Data}	$\frac{\Delta_{\text{Model}}}{\Delta_{\text{Data}}}$
$\iota - x_{\text{E}}$	incumbent innovation	25.53%	24.31%	5.02%	34.57%	14.52%
δ_d/δ	radical within external	8.46%	10.76%	-21.38%	-40.89%	52.29%
$\bar{\eta}/\lambda$	step-size ratio	34.49%	40.11%	-14.01%	-20.27%	69.12%
ι_s/ι	internal share	21.50%	24.22%	-11.23%	-22.95%	48.93%

Note: Δ_{Model} represents changes from the counterfactual to the model benchmark, Δ_{Data} is changes between the pre- and post-2008 period, both columns are presented in relative terms. Step-size ratio denotes the relative step-size of incremental to radical innovations.

Table D.7: Growth Decomposition with US Internal Patent Share

Δ_{Growth}	(i) quantity	(ii) quality-composition	(iii) quality-crowding	
-0.16	0.09	-0.04	-0.19	(p.p.)
	-56.25%	25.00%	118.75%	

Note: For each of the channels, we add the corresponding change in (i) ι; (ii) δ_d/δ and δ/ι; (iii) $\bar{\eta}$ to the pre-2008 economy, and see how it affects the aggregate growth rate. The second row shows the contribution of each channel, calculated by dividing the corresponding number by -0.16 p.p.

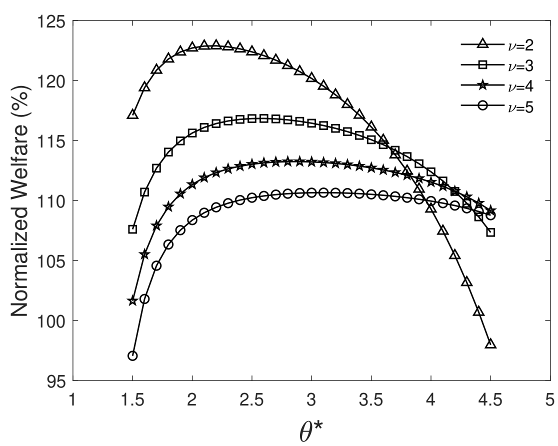
The weakening of all the growth-decomposition channels is more evident under the US calibration. As discussed, subsidies cause firms to pursue more external innovations, which pushes up wages and crowds out internal innovations in the general equilibrium. Hence the increase in total innovations ι is smaller than that in external ones δ, leading to a weakened quantity effect. As the decomposition formula shows, the importance of the quality effects hinges on the rate of creative destruction. Under the US calibration where internal patent share is high and creative destruction rate is relatively low, the quality-crowding effect, though still negative and dominant, inflicts less growth and welfare loss. However, even under the US calibration, the growth rate decline caused by subsidies

(-0.16 p.p.) is close to the baseline case (-0.19 p.p.), and the dominant force is still the *quality-crowding* channel. That again affirms the robustness of our baseline conclusions.

D.5 Social Optimum under Different Values of ν

We solve the planner's problem by using a brutal grid search on different values of θ^*. At each value, the supply of skilled and unskilled labor are determined as in Appendix B.1. We then let the demand side of the markets run until they all clear. The key difference between the planner's problem and the market equilibrium is that, skill premium in the planner's problem does not necessarily yield the θ^*_{SP} picked by the planner. We follow the literature and try $\nu \in [2, 5]$. Figure D.1 shows that social welfare is well hump-shaped w.r.t. θ^*. Moreover, the smaller ν is, the earlier welfare reaches its peak.

Figure D.1: Social Welfare as a Function of θ^*



Note: Under each value of ν, we solve the market equilibrium and normalize the corresponding welfare level to 100%.

D.6 Proof of Proposition 3

To implement θ^*_{SP}, we need to find combinations of (b_e, b_h) which solve

$$\frac{1-b_e}{\xi}\left[1-e^{-(r-g+d)}\right]\left(e^{-(r-g+d)}-(1-b_h)\frac{w^\ell}{w^h}\right)^{-1}=\theta^*_{SP}, \tag{13}$$

which can be rearranged to

$$b_h=-\frac{1-e^{-(r-g+d)}}{\xi\theta^*_{SP}}\frac{w^h}{w^\ell}b_e+\frac{1-(1+\xi\theta^*_{SP})e^{-(r-g+d)}}{\xi\theta^*_{SP}}\frac{w^h}{w^\ell}+1.$$

That is, policymakers face a linear trade-off between b_e and b_h when implementing θ^*_{SP}, while the slope of trade-off is endogenous.

To see that both b_e and b_h are necessary if θ^*_{SP} is low, we further examine the left-hand side (LHS) of equation (13). As shown in our quantitative analysis, the social planner's allocation is featured by high growth rate g and low skill premium w^h/w^ℓ. Moreover, the lower θ^*_{SP} is, the higher (lower) growth rate (skill premium) is. One can derive

$$\text{LHS}(b_e=0,b_h=1)=\frac{e^{(r-g+d)}-1}{\xi},$$

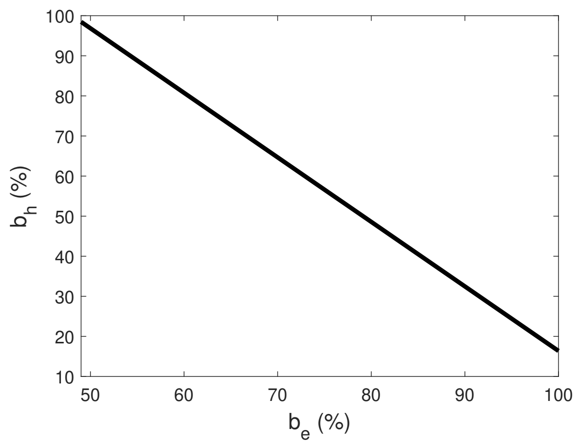
which is, the lowest θ^* a competitive equilibrium can reach without education subsidy. As θ^*_{SP} approaches 1, the difference between equilibrium interest rate and growth rate, $r-g=\rho+(\nu-1)\times g$, becomes greater, which eventually forces $\text{LHS}(b_e=0,b_h=1)$ to stay above the desired θ^*_{SP}. Thus, education subsidy is necessary if we want to implement a low θ^*_{SP}.

A necessary condition for equation (13) to hold is that its LHS stays positive, which in turn requires

$$\frac{w^h}{w^\ell}>(1-b_h)e^{(r-g+d)}.$$

Without skilled labor subsidy, i.e., $b_h=0$, this condition will be violated eventually since the skill premium approaches 1 when θ^*_{SP} approaches 1. Thus, skilled labor subsidy is also necessary if the θ^*_{SP} to be implemented is low. For example, to implement the socially optimal θ^*_{SP} in our benchmark case, policymakers can choose (b_e, b_h) located on the solid line in Figure D.2. Since the θ^*_{SP} is fairly low, policymakers need the workhorse of both subsidies to implement the desired allocation.

Figure D.2: Combinations of (b_e, b_h) to Implement $\theta^*_{SP} = 2.6$.



D.7 Subsidy Comparison

Here we compare the effects of various innovation subsidies in the model. More specifically, we have four kinds of subsidies: quantity-based subsidy b_n; generic R&D tax credit b_r; education subsidy b_e; and skilled labor subsidy b_h. We raise the magnitude of each subsidy by a small amount (5 percent) and check the changes in several important moments. Table D.8 summarizes the results.

The generic R&D tax credit, b_r, once strengthened, results in more R&D trials, and higher innovation quantity, but deteriorating innovation quality. Similar to the quantity-based subsidy, R&D tax credit is “quantity-biased” since it cannot distinguish between R&D

Table D.8: Effects from Strengthening the Subsidies

Variable	**Meaning**	**B.M.**	$\mathbf{b_n}$**+5%**	$\mathbf{b_r}$**+5%**	$\mathbf{b_e}$**+5%**	$\mathbf{b_h}$**+5%**
$R(x)/Vadd$	average R&D intensity	15.84%	16.02%	16.26%	15.73%	15.75%
δ_d/δ	radical share	8.01%	7.91%	7.83%	8.63%	8.51%
w^h/w^ℓ	skill premium	2.43	2.44	2.45	2.38	2.39
g	TFP growth rate	1.97%	1.96%	1.96%	2.07%	2.05%
U	social welfare	100%	99.80%	99.66%	101.14%	100.93%

Note: We strengthen each of the subsidies by 5 percent from the benchmark level while keeping others unchanged. The benchmark level of welfare is normalized to 100%.

expenditures on radical and incremental innovations. As a consequence, both subsidies contribute negatively to welfare if they were strengthened from their current levels.

Conversely, the two "quality-biased" subsidies, b_e and b_h, can effectively improve welfare by raising the skilled labor supply, reducing skill premium, and encouraging radical innovations. They effectively improve both the quantity and quality of innovations.