创业与管理学院

School of Entrepreneurship and Management

SHANGHAITECH SEM WORKING PAPER SERIES
No. 2022-003

Haste Makes Waste? Quantity-Based Subsidies under Heterogeneous Innovations

Linyi Cao

Institute for Advanced Research, Shanghai University of Finance and Economics

Helu Jiang

Institute for Advanced Research, Shanghai University of Finance and Economics

Guangwei Li

School of Entrepreneurship and Management, ShanghaiTech University

Lijun Zhu

Institute of New Structural Economics, Peking University

July 22, 2022

https://ssrn.com/abstract=4169354

School of Entrepreneurship and Management

ShanghaiTech University

http://sem.shanghaitech.edu.cn

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

Linyi Cao Helu Jiang Guangwei Li Lijun Zhu †

July 22, 2022

Abstract

With quantity-based innovation targets and subsidy programs launched since the mid-2000s, Chinese patents have seen a sharp increase in number, accounting for 46% of the world's total patent applications in 2020; however, their overall quality has been steadily declining over time. This paper develops a Schumpeterian growth model featuring innovating firms' quantity-quality trade-off between radical and incremental innovations, and decomposes subsidies' impact on growth and welfare into quantity and quality channels. We calibrate the model to Chinese firm-level R&D data in the early 2010s. The model-based quantitative analysis shows that the quality channel effects are negative and dominant, and quantity-based subsidies reduce welfare by around 9%. We evaluate welfare gains under a constrained planner's problem, and propose a quality-biased subsidy — subsidizing the human capital accumulation — which effectively recovers 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, Michael Song, Zhiwei Xu, Ge Ying, Bo Zhao and seminar and conference participants at SUFE Macro Workshop, Economic Fluctuation and Growth 2021 Fall Meeting, 2022 Asia Meeting of the Econometric Society, and 2022 Australasia Meeting of the Econometric Society.

†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

When economic growth first starts in a developing country, its source is typically investment in physical capital or capital-embodied technology diffusion from more developed countries. When the country gradually approaches the world technology frontier, a transition to innovation-driven growth is needed to achieve sustainable improvement in the residents' living standards. The evolution of the Chinese economy is one of the best examples to illustrate such dynamics. After three decades of rapid growth in the 1980s, 90s, and 2000s, arguably driven by factor accumulations, economic growth has significantly slowed down in China since the early 2010s. Partly due to the fear of falling into the so-called "middle-income trap", the Chinese government has launched a series of initiatives to ensure the country's success in transiting to an innovation-oriented economy since the mid-2000s (Ding and Li, 2015).

This paper empirically documents that, since the mid-2000s, the central and local Chinese governments have set quantity-based innovation targets in their five-year plan or other policy documentation. In particular, the total number of patents has been widely used as a concrete indicator of innovation achievement by Chinese governments. 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, which is 150% more than applications in the US, accounting for 45.69% of the global total.

Moreover, the number of patents per researcher in China progressed at a comparable rate to other advanced economies in the 1990s and early 2000s but increased much faster since the mid-2000s. By 2018, an average Chinese researcher produced patents almost twice as much as their US counterparts, raising concerns about the underlying patent quality. We further use information on a patent's forward citations or family size to measure its quality, and find that the overall quality of Chinese patents deteriorated steadily after the 2000s. In addition, we construct an innovation input-output dataset for the Chinese industrial firms, which enables us to understand the aggregate pattern from micro-level innovation decisions in the Chinese economy. We highlight the critical role of skilled labor in research and development (R&D) activities pursuing patents of high-quality.

Then, building on Schumpeterian growth models with heterogeneous innovations (Akcigit and Kerr, 2018; Acemoglu et al., 2018), we develop a general equilibrium framework featuring innovating firms' endogenous choices between radical and incremental innovations. Radical innovations significantly impact productivity, while incremental innovations build on existing radical innovations and make marginal improvements, with their impact gradually diminishing toward zero, capturing the worry that many patents may have minimal productive value in China's patent surge.

As R&D inputs, innovations use skilled and unskilled labor hired from competitive markets and entrepreneurial time. Each firm is endowed with 1 unit of non-tradable entrepreneurial time and decides how to allocate this scarce resource among two kinds of innovations. Motivated by empirical findings, 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. On the extensive margin, the economy admits two types of firms: high-type firms capable of pursuing radical and incremental innovations and low-type ones that only create incremental ones.

The government cannot precisely identify the quality of innovations and often bases its 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 encouraging overall innovations also bring an undesired shift of R&D efforts toward cheaper but low-quality 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. In the end, we decompose the impact of quantity-based subsidies on growth into three channels: 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 innovative industrial firms from 2011 to 2013. In particular, we use moments regarding high- and low-quality patents in data to discipline parameters related to radical and incremental innovations in the model, showing that the introduction of quantity-based innovation subsidies accounts for 42% of the quantity surge and 73% of the quality drop 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 0.33 percentage points, or 17% of the actual TFP growth decline from 2001-2007 to 2008-2014, and reduce the aggregate welfare by 9.39%.

China is still relatively scarce in innovative, skilled labor despite its fast economic catch-up. In 2018, 27% of the Chinese population between 25 and 34 years old have completed tertiary education, which is much lower than other major patent-holding economies. Within the model framework, we further propose and evaluate the impact of an alternative policy: skill subsidy or subsidizing the human capital accumulation, which effectively recovers the social planner's allocation. In the model skill is acquired through formal education before a worker enters the labor market. Skill subsidy reduces the cost of education and increases the skill supply. Since radical innovations are skill-intensive, increasing the supply of skilled labor substantially reduces the cost of R&D pursuing such inventions. Thus, in contrast to quantity-based innovation subsidies, skill subsidy is a quality-biased policy — it significantly promotes aggregate growth and welfare by raising 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 (Aghion and Howitt (1992), Grossman and Helpman (1991)) and its recent advancement with heterogeneous firms, e.g., Klette and Kortum (2004), Akcigit and Kerr (2018), and Acemoglu et al. (2020). Our model builds on Akcigit and Kerr (2018), who develop a model in which firms pursue radical and incremental innovations, and whether an R&D trial is radical or incremental is random. Unlike their work, we introduce a scarce R&D resource and allow the pursuit of different kinds of innovations to become an endogenous choice of the firm, generating a micro-level quantity-quality trade-off between high- and low-quality innovations. We also incorporate the heterogeneity of R&D input structure between high- and low-quality innovations and endogenize the human capital accumulation, a crucial dimension in understanding developing countries' innovation issues. We follow various works in the literature to discipline the model using patent data and study the aggregate implications of large-scale quantity-based subsidies adopted in China since the mid-2000s.

Our work also relates to studies investigating China's R&D policies and patent surges. Hu and Jefferson (2009) argue that amendments to the patent law, growth of FDI, and changing ownership are the major forces of the patent boom. Several studies have examined other hypotheses, including technological improvement, foreign direct investment, propatent legal changes, and the exit of low-patenting-propensity SOEs (Chen and Zhang, 2019; Fang et al., 2017; Ang et al., 2014; Jia and Ma, 2017). Li (2012) empirically studies the impact of innovation subsidy programs on patenting behaviors. The study confirms that local subsidy programs help stimulate patent applications. Using data from the China Employer-Employee Survey, Chen et al. (2019) find that innovation subsidies, which are more likely to be allocated to enterprises that are state-owned and with better political connections, positively impact incremental innovations but not radical ones.

A few recent papers study specific R&D policies in China using a more structural approach. Chen et al. (2021) study China's InnoCom program, which rewards a tax cut to firms with R&D investment above a certain threshold. They find that firms relabelling expenses as R&D accounts for a substantial fraction of reported R&D expenditures, and with relabelling, notch-based policies are more effective than tax credits. Wei et al. (2021) find that the InnoCom Program encourages rewarded firms to engage more in low-quality inventions and purchase patents from non-rewarded firms, and patent trading is a dominant channel for why the program hurts aggregate welfare. Neither paper emphasizes the quantity-quality trade-off facing innovating Chinese firms or studies innovation policies' impact on aggregate growth. Recently, König et al. (2020) develop a model featuring imitation and innovation decisions of firms subject to distortions to study the impact of R&D misallocation on long-term growth in China. They find that a large subsidy might even reduce the growth rate as it distorts firms' imitation-innovation decisions. The subsidy might also hurt growth in our framework but through a different channel.

Lastly, the paper is related to research on the role of human capital in innovation and economic growth. The idea that human capital affects the rate of technological change dates back to Nelson and Phelps (1966). Vandenbussche et al. (2006) develop a model in which productivity growth comes from either imitation or innovation. 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 basic empirical facts, and Section 3 introduces the model. The quantitative analysis is the focus of Section 4, and concluding remarks are presented in Section 5.

2 Institutional Background and Motivational Facts

This section first provides an overview of various quantity-based innovation targets set by China's central and local governments since the mid-2000s and associated innovation-promoting subsidies. Then, we empirically document the evolution of patenting behaviors in China, specifically, the patent surge and quality decline. Lastly, we illustrate the heterogeneity of input structure between firms that produce high-quality patents and those that do not, highlighting the role of skill in creating high-quality (radical) innovations.

2.1 Institutional Background

China's central and local governments have used quantity-based subsidy programs for a long time. 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., 2020). The general goal of science and technology development is to enhance independent innovation capability and turn China into an innovation-oriented country in 2020. One of the critical specific metrics 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. 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;

[1]Other specific targets listed in the documentation include the following. By 2020, the share of total R&D expenditure 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.

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 applied explicit targets on the number of patents. In Table 2.1, 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 western Guizhou province.

Table 2.1: 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 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.

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 Yaun (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]

2.2 Patent Surge

With the explicit quantity targets and associated subsidy policies, China has seen a dramatic surge in the total number of invention patents. However, facing quantity-based subsidies, firms may deviate from high-quality innovations to pursue easier goals — creating a larger number of low-quality patents. This section empirically documents that China's overall quality of patents has been steadily declining along with the quantity surge.

Sharp Increase in Quantity. There are three types of patents in China: invention, utility model, and industrial design. As applications of the last two do not require substantial review, we focus on invention patent, referred to as "patent" throughout the paper.[3] Figure 2.1 shows the evolution of the total number of patent applications in China and other major patent-holding economies.[4] The Chinese patent law was enacted in 1985. China's total number of patent applications in the 1980s and 90s was substantially smaller than Korea, Europe, Japan, and the United States. It surpassed Korea in the early 2000s and Japan and Europe later that decade. In 2011, China replaced the US as the world's No.1 patent application country. From 1985 to 2020, China's patent applications increased at an annual rate of 15.9%. It accounts for 1.02% of the world's patent applications in 1990, rising to 3.77% in 2000 and further to 19.58% in 2010. By 2020, this share increased to 45.69% of the world's total, which is larger than the combined share of patent applications from the US, Japan, Europe, and Korea, 44.87%, that same year.[5]

[2] Another widely adopted policy is the intellectual property rights pledge financing. In 2010, the Ministry of Finance issued the *Notice on Strengthening Intellectual Property Rights Pledge Financing to Support the Development of Small and Medium-Sized Enterprises*. Local governments have set up pledge financing schemes for intellectual property rights to subsidize small and medium-sized enterprises to reduce their borrowing costs for using intellectual property rights from pledge financing facilities. In Shenzhen city, the subsidized loan scheme provides a government subsidy of 40% on the total interest cost. By the end of 2012, 20 provinces have adopted pilot financial services on intellectual property rights pledge financing (Ding and Li, 2015).

[3] In 2020, China's total applications for invention, utility model, and industrial design patents were 1.50, 2.93, and 0.77 million, respectively. The latter two are not regarded as patents in some other countries, while invention patents are universally recognized.

[4] To accommodate more countries in the comparison, we use the number of patent applications in both Figures 2.1 and 2.2 instead of applied and eventually granted patents, as in the rest of the paper.

[5] Figure A.1 shows the number of granted patents from 1990-2020 in China and other economies. Note that there is a time lag between the application and grant of a patent; however, the general pattern in Figure A.1 is the same as in Figure 2.1.

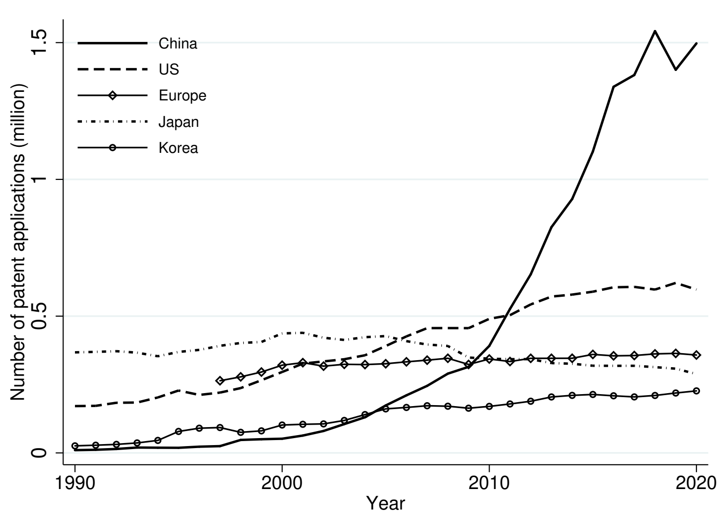


Figure 2.1: Evolution of Patent Quantity in China and Advanced Economies

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

Even though China's invention patents, as an innovation output, surpass that of other advanced economies, its progress in innovation inputs — an important one being the number of researchers — is less impressive. Table 2.2 shows the fraction of researchers among residents. Compared to other economies with strong patenting, China's innovation inputs are still relatively skill-scarce.

Figure 2.2 further compares the number of patents per researcher in China, and the US — China started at a much lower level at the beginning of the 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 decades, the Chinese government set quantity targets and adopted various innovation subsidies. As a result, 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.[6]

Note that the pattern in Figure 2.2 is not driven by differences in the two countries' patent grant rates, i.e., the fraction of applied patents in a given year that are eventually granted.

[6]There is a fallback in China's patents per researcher in 2019-2020. It is not clear if the drop is temporary or not without further data. Figure A.2 in the Appendix confirms that the pattern for other major advanced economies resembles that for the US.

Table 2.2: 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 labor force. Data source: USESCO.ORG.

Panel (a) of Figure A.3 shows the patent grant rates in both China and the US, indicating no substantial difference in either the level or the trend; panel (b) plots the number of applied and eventually granted patents per researcher, which still shows a rise in China's time series since the late 2000s.

The patent surge, in the total number or number per researcher, 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.

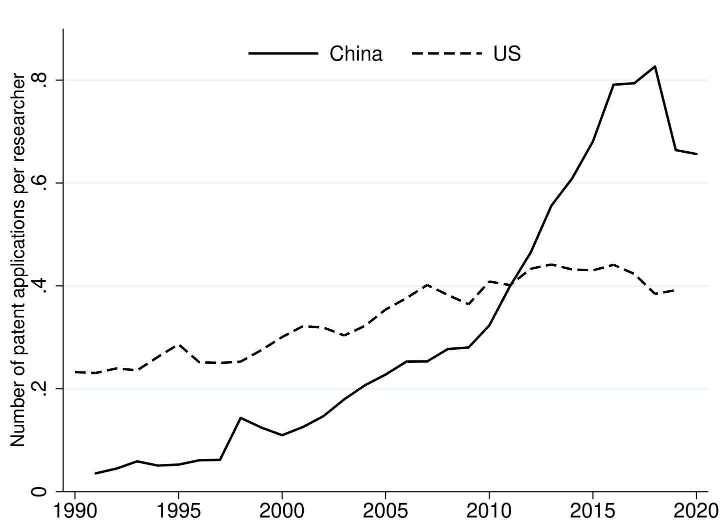


Figure 2.2: Number of Patent Applications per Researcher in China and the US

Note: Data source for Number of invention patent applications is World Intellectual Property Office, and for Number of researchers is OECD and China Statistical Yearbooks of Science and Technology.

Steady Decline in Quality. As the quantity of Chinese patents soars, their overall quality plunges. Since cross-country comparison is not needed for time series and micro-level patent data in China, we focus on applied and eventually granted patents as specified in Appendix A.1. In addition, we apply two criteria to measure the quality of a Chinese patent: (i) the fraction of forward citations a Chinese patent receives from US patents, and (ii) whether a Chinese patent's family size is greater than one, that is, has other countries' patents as its family members. We obtain the patent data in October 2020 and focus on patents applied before 2013 to avoid the truncation issue.

Figure 2.3 plots the evolution of the overall quality of Chinese patents based on the two measurements above. The share of patents with other countries' family members shows a steady decline in the 2000s and 2010s, and the share of forward citations from US patents among all posits a clear declining trend after the mid-2000s.[7]

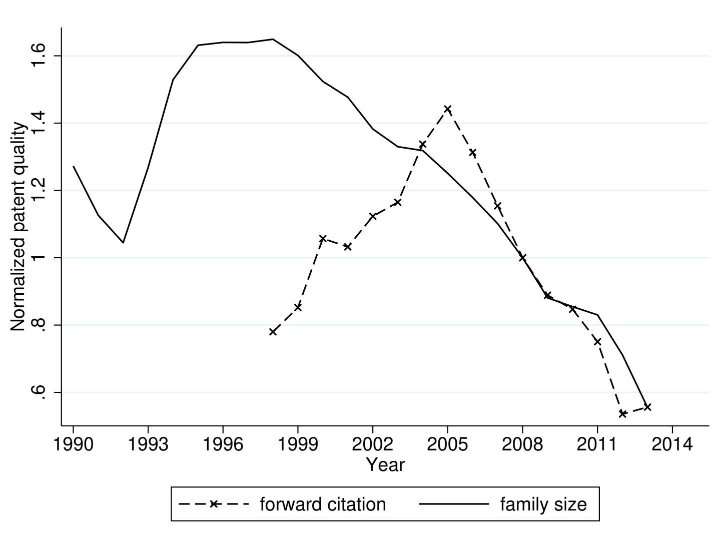


Figure 2.3: Evolution of Patent Quality

Note: This figure shows the share of forward citations from US patents divided by the average annual US forward citation rate, and the share of patents with other countries' family members among all Chinese patents. Values in 2008 are normalized to 1.

[7]Similar declining trend persists if the share of patents with US family members is used to measure overall quality. Wei et al. (2021) find a similar decline in Chinese patents' quality based on firm-level regression results.

2.3 Firm-Level Data: Skill Intensity and Firm Types

Our theoretical model hypothesizes that creating high-quality patents is more skill-intensive than making low-quality ones. Under such a hypothesis, a patent surge induced by quantity-based subsidies would increase demand for skilled labor and their wage, hurting efforts to pursue high-quality innovations and impairing overall patent quality. Due to the lack of information on inventors' skills at the patent level, this section constructs an input-output dataset on R&D activities of Chinese industrial firms and investigates the relation between skill intensity and patent quality at the firm level. We show that firms capable of creating high-quality patents tend to employ a larger share of skilled workers in R&D.

Data Source. We gathered information from three databases: (i) Annual Survey of Industrial Enterprises (ASIE), which covers all Chinese industrial firms with sales above 5 million RMB for the periods 1998-2013; (ii) a supplementary Firm Innovation Activity Database, available from 2008-2014, which contains industrial firms' R&D investments and skill composition of R&D personnel; (iii) Innography Patent Database, which provides information on patent ID, patent class, forward and backward citations, legal status, etc., from 1985 onwards. Combining ASIE, Innography, and Firm Innovation Activity Database, we construct an input-output dataset for firm-level R&D activities from 2008-2013.[8] We focus on applied and eventually granted patents and restrict to firms with records of at least one invention patent during the sample period. As the benchmark criteria, we label a Chinese patent as high-quality if any US patents have ever cited it.[9]

Table A.2 in the Appendix provides summary statistics of our firm-level final analysis sample.[10] From 2008 to 2013, innovating firms in our analysis sample account for 21.16% of total Chinese patents. The number of patents in the analysis sample increased at an annual rate of 19.3%. Meanwhile, the patent surge in the analysis sample is accompanied by a decrease in their overall quality. The share of high-quality patents in the analysis

[8]Note that patent citations are subject to time lags. More recent patents are less likely to be cited by others than older ones. We downloaded patent data from the Innography Database between October and December 2020, 7 years after the last year in our sample, 2013. To make a meaningful comparison, we examine five-year US citation rates and find a similar declining trend.

[9]We use this binary measure because (i) it is one of the mainstream measurements of quality in the literature; (ii) it is consistent with our theoretical framework, in which we model innovations in two ways (radical v.s. incremental); (iii) this measurement suffers little from the potential citation inflation issue due to the patent surge. We test for other definitions of patent quality measurements in Appendix C.2.

[10]Appendix A.1 illustrates the data sources, Appendix A.2 details variable construction. The data cleaning process and construction of the final sample can be found in Appendix A.3.

sample decreased dramatically, from 8% in 2008 to 2% in 2013.

As mentioned, due to a lack of information on skill input for individual patents, we rely on the firm-level skill composition of researchers to infer the skill intensity in producing different kinds of innovations. In particular, we identify a firm as high-type if it creates at least one high-quality patent over the sample period. Among an innovating firm's R&D personnel, we label those with a medium or senior professional title (*zhonggaoji zhicheng*) as skilled labor. Skill intensity is defined as the ratio between skilled labor and total R&D personnel (*keji huodong renyuan*). Table 2.3 shows the skill composition and firm type distribution over time. From 2011 to 2013, the skill intensity was 31.30% for high-type firms and 24.52% for low-type firms, while high-type firms accounted for 22.34% of total firms.[11]

Table 2.3: Skill Intensity and Firm Types

	2011	2012	2013	Average
Skill intensity of high-type firms	32.10%	31.56%	30.32%	31.30%
Skill intensity of low-type firms	26.43%	24.38%	23.36%	24.52%
Fraction of high-type firms	26.14%	22.54%	19.71%	22.34%

Note: This table shows the skill intensity (top panel) and fraction of high-type firms (bottom panel) over time.

Firms producing high-quality patents tend to employ a higher fraction of skilled labor among their R&D personnel, which we take as evidence that producing high-quality patents is more skill-intensive. We further verify this statement using a regression, as shown in Table A.3. Controlling for firm characteristics including firm age, (log) employment, (log) revenue, and (log) assets as well as industry, year, location, and ownership type fixed effects, column (3) of the table shows that the firm's skill intensity positively correlates with its patent quality, measured by the share of the firm's high-quality patents. This reinforces our hypothesis that high-quality innovations require proportionately more skilled labor.

Facing quantity-based subsidies, firms may find it more profitable to maximize the number of innovations to obtain maximum subsidy. Such distorted incentives imply a potential deviation from high-quality innovations that require more skilled labor input to

[11] Skill composition data is not available for 2008 and 2009.

cheaper but low-quality ones. A micro-level quantity-quality trade-off facing innovating firms and quantity targets may help explain the observed patent surge along with their quality plunge. In that spirit, we build a Schumpeterian growth model featuring heterogeneous firms and innovations to evaluate the impact of quantity-based innovation policies on economic growth and highlight the benefit of quality-biased innovation-promoting policies, such as subsidizing the human capital accumulation.

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 who 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 used in consumption and entrant firms' overhead investment, which we specify below. The final good is 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 $\omega \in [0,1]$ at time t, and $q_\omega(t)$ denotes its quality. The parameter $\epsilon \in (0,1)$ governs the value added share of intermediate varieties.

$N(t)$ is the number of packagers. The total supply of packagers from the household is assumed to be exogenously fixed at 1, so their competitive wage is $w^N = \epsilon Y(t)$. In the end, the final good producers' demand for intermediate variety ω is simply given by

$$p_\omega = q_\omega^{\epsilon} y_\omega^{-\epsilon}. \tag{3}$$

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, providing the highest quality

q_ω.[12] Denote $\mathcal{F}$ the total measure of incumbent firms in the economy. 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 productivity portfolio. Denote n_f the cardinality of the set Q_f, which represents the number of product lines that the firm owns. A firm that loses all product lines exits the economy permanently, so we have $n \geq 1$ for incumbent firms.

Production of intermediate goods requires unskilled labor as the sole input and takes the following form

$$y_\omega(t) = \bar{q}(t)\ell_\omega(t), \tag{4}$$

where $\bar{q}(t) \equiv \int_0^1 q_\omega(t)d\omega$ is the economy-wide average productivity at time t, which captures the cross-firm "spillover" effect of innovations.

To maintain simplicity and avoid limit pricing, we follow the standard approach in the Schumpeterian growth literature and assume a two-stage price-bidding game (Acemoglu et al., 2012).[13] In equilibrium, the firm owning the leading technology can charge a monopolistic price until being replaced in the future by a successful innovator. Given this setting, the profit maximization problem of the firm that owns the leading technology in product line ω is

$$\max_{y_\omega} \quad q_\omega^\epsilon y_\omega^{1-\epsilon} - \frac{w^\ell}{\bar{q}} y_\omega. \tag{5}$$

It follows that $p_\omega = \frac{1}{1-\epsilon}\frac{w^\ell}{\bar{q}}$. That is, firms will charge a constant markup $\frac{1}{1-\epsilon}$. The profit from owning product line ω then is

$$\pi_\omega = \epsilon \left[(1-\epsilon)\frac{\bar{q}}{w^\ell}\right]^{\frac{1-\epsilon}{\epsilon}} q_\omega \equiv \pi q_\omega. \tag{6}$$

Therefore, profit is a linear function of the product line's quality if $w^\ell/\bar{q}$ is a constant, which is true on the balanced growth path.[14]

R&D Heterogeneity. In addition to production, intermediate-goods firms also spend on R&D to pursue innovations. Following the literature, R&D efforts are assumed to be

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

[13] 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.

[14] The labor demand is also linear in q_ω: $\ell_\omega = \left[(1-\epsilon)\frac{\bar{q}}{w^\ell}\right]^{\frac{1}{\epsilon}} \frac{q_\omega}{\bar{q}}$.

undirected. Upon a successful innovation, a firm improves the quality of a random product line by a step size from its current frontier. Innovations are heterogeneous; specifically, two kinds of innovation exist: radical vs. incremental. Incremental innovations build on one existing radical innovation. As in Akcigit and Kerr (2018), we assume the quality improvement associated with radical innovation is fixed and large, while that of incremental innovation is small and gradually diminishes toward zero.

R&D uses skilled labor, unskilled labor, and entrepreneurial time as inputs.[15] Each firm is endowed with 1 unit of non-tradable entrepreneurial time. When a firm owning n product lines hires h units of skilled labor, ℓ units of unskilled labor, and allocates e fraction of entrepreneurial time to pursue *radical* innovations, it adds one more product line to its portfolio at the following Poisson flow rate

$$X_d = z_d \, n^{1-\phi} \left(e^{\sigma} h^{\gamma_d} \ell^{1-\gamma_d} \right)^{\phi}, \tag{7}$$

where $z_d \geq 0$ is the firm's productivity in pursuing radical innovations. Parameter $\phi \in (0,1)$ is the elasticity of successful innovation concerning R&D.[16] The parameter $\gamma_d \in (0,1)$ is the skill intensity of radical innovation. The parameter $\sigma > 0$ denotes the elasticity of innovation arrival rate on entrepreneurial time. When $\sigma \to 0$, we return to a typical R&D function used in the literature. As the total endowed entrepreneurial time is fixed at 1, a positive value of σ allows us to examine individual firm's trade-off between different kinds of innovations, a point we illustrate in detail below.

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

$$q_{\omega}(t_+) = q_{\omega}(t) + \lambda \bar{q}(t), \tag{8}$$

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

Firms can also pursue a second incremental innovation. When a firm owning n product lines hires h units of skilled labor, ℓ units of unskilled labor, and allocates the remaining $1 - e$ units of entrepreneurial time in incremental innovations, it adds one more product

[15] We introduce "entrepreneurial time" to capture R&D inputs that are scarce and non-tradable, such as a manager's time to supervise R&D projects, or new ideas from a research team, etc.

[16] It also captures a within-firm "spillover" effect from existing innovations to the creation of a new one.

line to its portfolio at the following Poisson flow rate

$$X_m = z_m \, n^{1-\phi} \left((1-e)^{\sigma} h^{\gamma_m} \ell^{1-\gamma_m} \right)^{\phi}, \tag{9}$$

where z_m and γ_m are different from that of the radical innovations. Following the empirical result in Section 2.3, we assume $\gamma_m < \gamma_d$; that is, incremental innovations are less skill-intensive than radical ones.

The quality improvement following a successful incremental innovation depends on the distance from the most recent radical innovation, i.e., times of incremental improvement already created in the 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 step-size of quality improvement would be

$$q_\omega(t_+) = q_\omega(t) + \eta \alpha^{\tau_\omega - 1} \bar{q}(t), \tag{10}$$

where $\eta \in (0, \lambda)$ governs the initial step-size, and $\alpha \in (0, 1)$ governs how fast the effect diminishes. The idea behind this setting is that the effect of incremental innovations weakens until a radical one arrives and resets the clock.

Now we are ready to derive the associated R&D cost function. 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,[17] and w^h and w^ℓ wage rates for skilled and unskilled workers, respectively.

For an individual firm whose innovation intensities are (x_d, x_m), the associated function of R&D cost per line, $R(x_d, x_m)$, is given by[18]

$$R(x_d, x_m) = \left[\Theta_d (x_d)^{\frac{1}{\sigma+1}} + \Theta_m (x_m)^{\frac{1}{\sigma+1}} \right]^{\sigma+1}, \tag{11}$$

[17] Our specification yields innovation intensity of $x_i = z_i e_i^{\sigma\phi} \left(h^{\gamma_i} \ell^{1-\gamma_i}/n \right)^{\phi}$, for $i = d, m$. The term $z_i e_i^{\sigma\phi}$ represents the "effective productivity" when e_i amount of entrepreneurial time is allocated.

[18] The cost function is derived from a standard cost minimization problem. Minimize the total R&D spending on skilled and unskilled labor, subject to a given output of innovation intensities x_d and x_m from the previously specified production functions.

where

$$\Theta_i(x_i) \equiv \Delta_i \left(w^h\right)^{\gamma_i} \left(w^\ell\right)^{1-\gamma_i} \left(\frac{x_i}{z_i}\right)^{\frac{1}{\phi}},$$

and

$$\Delta_i \equiv \gamma_i^{-\gamma_i}(1-\gamma_i)^{\gamma_i - 1}, \quad \text{for } i = d, m.$$

The cost function has two components, Θ_d and Θ_m, each associated with labor inputs in corresponding R&D activities. These two components add up nonlinearly because of the fixed and non-tradable entrepreneurial time each firm has as an endowment. We assume $\sigma < (1-\phi)/\phi$ to maintain decreasing return to scale and avoid corner solutions in x_d or x_m. This argument shall become clearer when we arrive at the incumbent firm's value function. Also, note that when $\sigma \to 0$, the economy approaches the case in which choices of x_d and x_m are separable.

Quantity-Based Subsidy. Though radical and incremental 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 high-quality patents and incremental innovations to low-quality ones.[19] Total number of active patents that a firm has is therefore proportional to the number of product lines that the firm controls.

We define quantity-based subsidy to innovating firms as any subsidy that rewards the number of active patents a firm holds, i.e., n, disregarding the underlying quality. In particular, we use the form $n \times b_n \pi \bar{q}$, where b_n denotes the subsidy-to-profit ratio. 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 subsidies — cheaper land cost, accessibility to loans, etc., that an innovating firm receives as long as the subsidies are quantity-based.

Firm Heterogeneity. 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, that is, $z_{\mathrm{H}d} > 0$ and $z_{\mathrm{H}m} > 0$. The low-type (L) are capable of pursuing only incremental ones, that is, $z_{\mathrm{L}m} > 0$ but $z_{\mathrm{L}d} = 0$.[20]

[19] The number of patents an innovation embodies might change over time.

[20] This distinction between firms allows us (i) to infer R&D input structure from observable firm-level data, and (ii) to investigate both the intensive margin (changes in the intensity of different kinds of innovations within high-type firms) and the extensive margin (changes in the share of high- and low-type firms in the economy) associated with innovation policies.

At any point, there are a total mass 1 of potential entrants. Upon a successful innovation, the potential entrant enters the economy with one product line in its portfolio. Potential entrants are of low-type by default; however, after paying an overhead investment of $K(p)$, they receive a probability $p \in [0, 1]$ turning into high-type entrants. For tractability, we assume that entrants stick to their chosen type throughout the life cycle. Section 3.1 provides a more detailed discussion about the entrant firms' problem and the incumbent firms' type and product line distribution.

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 joins the economy, so the total population of workers stays constant. Upon entry, each individual randomly draws a type θ from a *Pareto* distribution of talent[21]

$$\mathbb{P}\left\{\theta \leq \tilde{\theta}\right\} = 1 - \tilde{\theta}^{-2}, \text{ for } \tilde{\theta} \in [1, \infty).$$

An individual can work as an unskilled worker without any investment in education; however, they must spend time in school to become skilled.[22]

An individual's cost of education is negatively associated with their talent type. In particular, it requires $1/\theta$ units of education service for an individual of type θ to become skilled labor. Education service is produced by existing skilled workers employed in the education sector at a competitive wage rate and with technology

$$e = \xi \, h^{\text{teacher}}, \tag{12}$$

where h^{teacher} denotes the mass of skilled workers employed in education. We introduce the productivity in the education sector, $\xi \in (0, \infty)$, to capture the overall efficiency of an economy's education infrastructures.

Young people choose to invest in education and become skilled if and only if the expected lifetime return from doing so — earning a skilled wage minus paying the education cost — surpasses the lifetime value of being an unskilled worker earning a lower wage. Appendix B.1 shows that a young individual chooses to invest in education if and only if

[21] As become clearer in Section 4, the Pareto distribution's shape parameter is a free parameter. We simply fix it at 2, the infimum for a finite variance.

[22] The way we introduce education and endogenous human capital accumulation follows Acemoglu et al. (2018).

her type is above a certain threshold θ^*, which depends on the economy's skill premium w^h/w^ℓ and the productivity in the education sector ξ. Lastly, we assume that producing education services require a fixed amount of unskilled labor, $\ell^{\text{edu}} > 0$, to properly match the size of the education sector in the quantitative exercise.

3.1 Equilibrium

We focus on equilibrium featuring a balanced growth path; that is, the average productivity of the economy, $\bar{q}(t)$, grows at a constant rate g, while other aggregate variables grow proportionally.

Incumbent Firm's Value Function. The state variables of an incumbent firm include its type, product portfolio Q, and the economy's average productivity $\bar{q}$. Denote r the interest rate and δ the endogenous creative destruction rate in the economy. Both rates are determined on the aggregate level and taken as given for an individual firm.

The value function for a high-type firm is written as

$$\begin{aligned} rV_{\mathrm{H}}(Q,\bar{q}) - \dot{V}_{\mathrm{H}}(Q,\bar{q}) = \max_{x_d,x_m} & \sum_{q_\omega \in Q} \Big\{ \underbrace{\pi q_\omega}_{\text{profit}} + \underbrace{\delta \big[V_{\mathrm{H}}\left(Q \backslash \{q_\omega\}, \bar{q}\right) - V_{\mathrm{H}}(Q,\bar{q}) \big]}_{\text{loss from creative destruction}} \Big\} \\ & + \underbrace{n \times x_d \Big[\mathbb{E}_{\omega'} V_{\mathrm{H}}\left(Q \cup \{q_{\omega'} + \lambda \bar{q}\}, \bar{q}\right) - V_{\mathrm{H}}(Q,\bar{q}) \Big]}_{\text{return from radical innovations}} \\ & + \underbrace{n \times x_m \Big[\mathbb{E}_{\omega'} V_{\mathrm{H}}\big(Q \cup \{q_{\omega'} + \eta \alpha^{\tau_{\omega'}-1} \bar{q}\}, \bar{q}\big) - V_{\mathrm{H}}(Q,\bar{q}) \Big]}_{\text{return from incremental innovations}} \\ & - \underbrace{n \times R(x_d, x_m)}_{\text{R\&D cost}} + \underbrace{n \times b_n \pi \bar{q}}_{\text{quantity-based subsidy}} \quad . \end{aligned} \tag{13}$$

The first line is static profit from each product line, plus the value change due to creative destruction. $Q\backslash\{q_\omega\}$ denotes the remaining portfolio of the firm after losing line ω due to a successful innovation by another firm. The second line is the net value change from a successful radical innovation, which adds a random product line ω' into the firm's product portfolio. The expectation is over ω' as which line the innovation lands on is random. As described before, a successful radical innovation raises the product quality by $\lambda\bar{q}$. The third line is the net value change following a successful incremental innovation. The last line includes the R&D cost and all quantity-based subsidies to the firm. Total subsidies a firm obtains equals its number of valid patents, n, times the subsidy for each patent, $b_n\pi\bar{q}$, where b_n denotes the subsidy-to-profit ratio, and $\bar{q}$ is added to make sure the existence of a balanced growth path.

Similarly, we can write the value function of a low-type firm as

$$\begin{aligned} rV_{\mathrm{L}}(Q,\bar{q}) - \dot{V}_{\mathrm{L}}(Q,\bar{q}) = \max_{x_m} \sum_{q_\omega \in Q} & \left\{ \pi q_\omega + \delta \left[V_{\mathrm{L}}\left(Q \backslash \{q_\omega\}, \bar{q}\right) - V_{\mathrm{L}}(Q,\bar{q}) \right] \right\} \\ & + n \times x_m \left[\mathbb{E}_{\omega'} V_{\mathrm{L}}\left(Q \cup \{q_{\omega'} + \eta \alpha^{\tau_{\omega'}-1} \bar{q}\}, \bar{q}\right) - V_{\mathrm{L}}(Q,\bar{q}) \right] \\ & - n \times \Theta_m(x_m) + n \times b_n \pi \bar{q}. \end{aligned} \tag{14}$$

The value of a low-type firm resembles that of a high-type firm, except that low-type firms are unable to pursue radical innovations since we assumed that $z_{\mathrm{L}d} = 0$.

One important equilibrium property of the optimization problems is that firms of the same type choose the same innovation intensity per line, regardless of their differences in product lines n or product portfolio Q. The property that product lines of a firm are "separable" in this class of models traces back to Klette and Kortum (2004). We end up tracking three innovation intensities in equilibrium: $x_{\mathrm{H}d}$, $x_{\mathrm{H}m}$ for high-type firms, and $x_{\mathrm{L}m}$ for low-type firms. The same logic applies to R&D spending per line, the value of the firm per line, etc. Appendix B.4 provides a derivation of the linearity of $V_i(Q,\bar{q})$ in terms of n.

Entrant's Value Function. As previously explained, there is a total mass 1 of potential entrants, who are pursuing only incremental innovations at a fixed Poisson rate $x_{\mathrm{E}} > 0$. Entrants are of low-type by default; however, they can become high-type with probability $p \in [0,1]$ by making a one-time overhead investment

$$K(p) = \left[-\ln(1-p) - p \right] \chi \bar{q}, \tag{15}$$

where $\chi > 0$ is a cost coefficient. The overhead investment uses the final good.[23]

The value function for a potential entrant can be written as

$$rV_{\mathrm{E}} = x_{\mathrm{E}} \left[\max_{p} \left\{ pV_{\mathrm{H}} + (1-p)V_{\mathrm{L}} - K(p) \right\} - V_{\mathrm{E}} \right], \tag{16}$$

[23] The setup captures that, firms or institutions need to make investments, such as building infrastructures or laboratories, to pursue high-quality and cutting-edge innovations.

where

$$V_{\rm i} \equiv \mathbb{E}_{\omega'} V_{\rm i} \left(\{q_{\omega'} + \eta \alpha^{\tau_{\omega'}-1} \bar{q}, \}, \bar{q} \right), \quad i = H, L$$

are the expected values of an i-type firm with one product line.

Since all entrant firms are ex-ante identical, they end up choosing the same overhead investment $K(p^*)$. In equilibrium, the fraction of high-type entrants, p^*, is given by

$$p^* = \frac{(V_{\rm H} - V_{\rm L})}{(V_{\rm H} - V_{\rm L}) + \chi \bar{q}}.$$

Stationary Distribution. We focus on a balanced growth path where aggregate variables grow at a constant rate $g \equiv \dot{\bar{q}}(t)/\bar{q}(t)$, and relevant distributions are stationary. Note that the model's distribution of quality across product lines does not matter as long as its mean $\bar{q}(t)$ exists and grows at a constant rate. We need to monitor two distributions, one over τ, i.e., the step-size of incremental innovations, and the other over n, i.e., the number of product lines.

Denote δ_d and δ_m the creative destruction rate due to radical and incremental innovations, respectively. Further denote δ the aggregate creative destruction rate, that is, $\delta \equiv \delta_d + \delta_m$. As shown in Appendix B.2, the expected step-size of an incremental innovation is given by

$$\bar{\eta} = \eta \bigg/ \left(\alpha + \frac{1-\alpha}{\delta_d/\delta} \right). \tag{17}$$

The expected quality improvement from incremental innovations decreases with a faster decay rate, i.e., a smaller α. Additionally, as the fraction of incremental innovations on the aggregate level δ_m/δ increases, the expected step-size also becomes smaller. This property captures the positive (negative) externality of radical (incremental) innovations.

Denote $\mu_{j,n}$ the mass of j-type firms who owns n product lines, where j = H,L. The expression of $\mu_{j,n}$ under a stationary distribution is derived in Appendix B.2. We can write the creative destruction due to radical and incremental innovations as[24]

$$\delta_d = \sum_{n=1}^{\infty} \mu_{{\rm H},n} \times n x_{{\rm H}d}; \quad \delta_m = \sum_j \sum_{n=1}^{\infty} \mu_{j,n} \times n x_{jm} + x_{\rm E}. \tag{18}$$

[24] Recall that only high-type firms can pursue radical innovations, and all potential entrants are assumed to be pursuing incremental innovations.

And the economy-wide creative destruction rate, δ, is

$$\delta = \sum_j \sum_{n=1}^{\infty} \mu_{j,n} \times n x_j + x_{\mathrm{E}}, \tag{19}$$

where $x_{\mathrm{H}} = x_{\mathrm{H}d} + x_{\mathrm{H}m}$ and $x_{\mathrm{L}} = x_{\mathrm{L}m}$ are total innovation intensities per line for each type of firms.

Moreover, the total number of product lines always sums to 1. Formally, we give the following proposition, whose proof can be found in Appendix B.3

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

The model also generates interesting implications regarding product lines and product quality. The average number of product lines owned by the j-type firm is given by $\bar{n}_j \equiv \left(\sum_n \mu_{j,n} \times n \right) \Big/ \sum_n \mu_{j,n}$. Based the Appendix B.2 results, firms with higher innovation intensity shall own more product lines in expectation.

Furthermore, firms of different types are heterogeneous in average product quality, so as average profit or revenue, which are linear dependents of the product quality. A general property is that firms pursuing a higher ratio of radical innovations, x_{jd}/x_j, enjoy a higher average product quality. Like our Proposition 1, both properties can be extended to $N \geq 2$ many types of firms; we revisit them in our quantitative analysis in Section 4.

Market Clearing. Market clearing conditions for the aggregate economy are straightforward. For unskilled labor, it is

$$\sum_j \sum_{n=1}^{\infty} \mu_{j,n} \times n \hat{\ell}_j + \int_0^1 \ell_\omega d\omega + \ell^{\mathrm{edu}} = \ell^{\mathrm{supply}}, \tag{20}$$

where $\hat{\ell}_j$ denotes the unskilled labor employed per line in R&D activities of the j-th type incumbent firms, where again $j = $ H,L in our baseline economy. ℓ_ω denotes the unskilled labor employed in producing intermediate variety ω, and ℓ^{edu} denotes that employed in the education sector, which is exogenous.

Correspondingly, the market clearing condition for skilled labor is

$$h^{\text{R\&D}} + h^{\text{teacher}} = h^{\text{supply}}, \tag{21}$$

where

$$h^{\text{R\&D}} = \sum_j \sum_{n=1}^{\infty} \mu_{j,n} \times n\hat{h}_j.$$

Again $\hat{h}_j$ denotes the skilled labor employed per line in R&D activities of incumbent firms of the j-type, and $h^{\text{R\&D}}$ is the total mass of skilled workers employed in the R&D sector. Here innovation intensity per line x_j is identical within each type, thus, so are $\hat{h}_j$ and $\hat{\ell}_j$. Detailed derivations on the supply side can be found in B.1.

Lastly, the market clearing condition for the final good is

$$C + K(p^*) \times x_{\text{E}} = Y. \tag{22}$$

3.2 Properties of the Economy

3.2.1 The Quantity-Quality Trade-off

Firms in the economy face a quantity-quality trade-off between creating more innovations (i.e., a larger x) and creating better innovations (i.e., a higher ratio of x_d/x). Innovation subsidies may impact such trade-offs, as seen from an individual firm's optimal decisions. Appendix B.4 indicates that the value function can be expressed as $V(Q, \bar{q}) = \sum_i Aq_i + nB(b_n)\bar{q}$. The first term denotes profit from owning product lines, while the second term contains net values from R&D, which depends on innovation subsidies b_n. Regarding firms' choices over innovation, we have the following proposition.

Proposition 2. *The ratio between radical and incremental innovation intensities of a high-type firm satisfies the following condition*

$$\left[\frac{x_d}{x_m}\right]^{\frac{1}{\phi(\sigma+1)}-1} \propto \underbrace{\frac{A\lambda + B(b_n)}{A\bar{\eta} + B(b_n)}}_{\text{innovation return}} \times \underbrace{\left[\frac{w^h}{w^\ell}\right]^{-(\gamma_d-\gamma_m)}}_{\text{input structure}} \times \underbrace{\left[\frac{z_d}{z_m}\right]^{\frac{1}{\phi(\sigma+1)}}}_{\text{R\&D productivity}}. \tag{23}$$

We assumed $\sigma < (1-\phi)/\phi$, so the power coefficient on the left-hand side is positive. The first term on the right-hand side (RHS) captures the ratio of returns from radical to incremental innovations. The direct return of radical innovation is from its productivity improvement effect, as captured by $A\lambda$. Similarly, that of incremental innovation is captured by $A\bar{\eta}$. The fact $\lambda > \bar{\eta}$ indicates that the direct return of radical innovations is greater.

The indirect return, $B(b_n)$, is primarily affected by (quantity-based) innovation subsidies and is identical for both innovations. A sizable generic subsidy shrinks the difference in total returns from radical and incremental innovations, which raises firms' incentive to pursue proportionately more incremental innovations. Furthermore, $\gamma_d > \gamma_m$, that is, radical innovations are more skill-intensive than incremental ones. If all firms are incentivized to do more research, which raises the equilibrium skill premium, w^h/w^ℓ, they optimally allocate more R&D efforts to incremental innovations.

Lastly, the elasticity of innovation creation on entrepreneurial time, σ, primarily affects the degree of quantity-quality trade-off facing innovating firms. When σ is larger — not exceeding its upper bound — the trade-off gets more substantial, and we expect more significant responses from innovative firms to quantity-based subsidies. Moreover, the impact magnifies drastically as σ approaches the upper bound. To see that, we take a log on both sides of the equation (23) and get

$$\log\left(\frac{x_d}{x_m}\right) = \underbrace{\frac{\phi(\sigma+1)}{1-\phi(\sigma+1)}}_{\text{impact coefficient}} \times \log\left(\text{RHS}\right). \tag{24}$$

As σ approaches $(1-\phi)/\phi$, the impact coefficient, which represents the magnitude of responses from innovating firms, rises to ∞ at an accelerating rate. This property is helpful for us to identify σ in the forthcoming quantitative analysis.

Quality Change. The aggregate share of radical innovations, δ_d/δ, equals to that share within high-type firms, $x_{\mathrm{H}d}/x_{\mathrm{H}}$, times the fraction of innovations created by high-type firms, $\delta_{\mathrm{H}}/\delta$.[25] It follows that changes in innovation quality can be expressed as

$$\Delta\frac{\delta_d}{\delta} = \underbrace{\Delta\frac{x_{\mathrm{H}d}}{x_{\mathrm{H}}} \times \frac{\delta_{\mathrm{H}}}{\delta}}_{\text{intensive margin}} + \underbrace{\frac{x_{\mathrm{H}d}}{x_{\mathrm{H}}} \times \Delta\frac{\delta_{\mathrm{H}}}{\delta}}_{\text{extensive margin}}. \tag{25}$$

[25] Consistent with our definitions of δ_d and δ_m, we define $\delta_{\mathrm{H}} \equiv \sum_{n=1}^{\infty} \mu_{\mathrm{H},n} \times n x_{\mathrm{H}}$.

The intensive margin comes from the quantity-quality trade-off facing high-type firms, while the extensive margin is attributed to entrant firms' endogenous type choice, p^*.

3.2.2 Aggregate Growth and Welfare

Along a balanced growth path, the aggregate welfare is

$$U = \int_0^\infty \exp(-\rho t)\,\frac{C(t)^{1-\nu}-1}{1-\nu}\,dt = \frac{1}{1-\nu}\left[\frac{{C_0}^{1-\nu}}{\rho-(1-\nu)g} - \frac{1}{\rho}\right]. \tag{26}$$

A critical determinant of the welfare is the aggregate growth rate, g, which satisfies

$$g = \delta_d\lambda + \delta_m\bar{\eta} = \delta\left[\frac{\delta_d}{\delta}\lambda + \left(1-\frac{\delta_d}{\delta}\right)\bar{\eta}\right], \tag{27}$$

where $\bar{\eta}$ denotes the expected step-size of incremental innovations. As δ_d and δ_m are aggregate quantity of radical and incremental innovations, the growth rate can be viewed as a weighted sum of their step-sizes. The growth rate differential, e.g. between economies with and without a particular policy, can be decomposed into the following terms

$$\Delta g = \underbrace{\Delta\delta\times\left[\frac{\delta_d}{\delta}\lambda + \left(1-\frac{\delta_d}{\delta}\right)\bar{\eta}\right]}_{\text{(i) quantity-creative destruction}} + \underbrace{\delta\times\left[\Delta\frac{\delta_d}{\delta}\times(\lambda-\bar{\eta})\right]}_{\text{(ii) quality-composition}} + \underbrace{\delta\times\left[\left(1-\frac{\delta_d}{\delta}\right)\times\Delta\bar{\eta}\right]}_{\text{(iii) quality-crowding}}. \tag{28}$$

The first term refers to the quantity channel, while the second and third are the quality channels. The first "quantity-creative destruction" term captures that the aggregate growth rate changes if a policy induces a change in the aggregate creative destruction rate, δ, or equivalently the total number of newly created patents. A policy that changes aggregate innovation quality, δ_d/δ, further affects aggregate growth. (a) It changes the composition of radical and incremental innovations, which generate different productivity impacts, captured by the second "quality-composition" channel. (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.

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 aggregate- and firm-level data, and evaluate the impact of *quantity-based* subsidies on patent quantity surge, quality drop, 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 human capital accumulation, which we show effectively recovers the planner's allocation.

4.1 Calibration Strategy

We calibrate the model's benchmark economy to 2011-2013 firm-level data. Before detailing the calibration strategy, we first elaborate on how we identify the parameter of σ, which governs the micro-level quantity-quality trade-off facing innovating firms, from firms' responses to policy variations.

4.1.1 High-Tech Enterprises and Identification of σ

To identify σ, we need a policy that generates heterogeneity among high-type firms. The logic behind resides in Proposition 2 and equation (24). The policy variation we utilize is the recognition of High-Tech Enterprises (HTE) under the Chinese InnoCom Program. More specifically, we utilize the fact that firms recognized as HTEs can enjoy extra returns from their patents. Appendix C.1 provides a brief introduction to the InnoCom Program and the recognition of HTEs. Correspondingly, we extend the baseline model to allow for four types of firms: high-type & HTE; high-type & non-HTE; low-type & HTE; low-type & non-HTE. HTEs and non-HTEs differ in that the former receives an extra subsidy, Δb_n, on top of the uniform quantity subsidy b_n.

Table C.1 shows that from 2011 to 2013, around 55% of entrant firms in our sample are HTEs. The difference in the probability of HTE recognition between high- and low-type entrants is negligible: 53.4% for the former and 55.6% for the latter. Therefore, we assign all entrants the same exogenous probability of being entitled HTE, $p^{\text{HTE}} = 55\%$.[26] For model tractability, we further assume that HTE or non-HTE status is permanent throughout a firm's life cycle.[27]

[26] The minimum requirement for applying to be HTE is 1 invention patent, so the incentive for firms to accumulate invention patents before qualification is non-substantial.

[27] In reality, the HTE title needs to be renewed every 3 years; however, the assumption of permanent status is justified on the following grounds. First, the probability of successive renewal is high. Second, this HTE assignment setup generates an incumbents' type distribution very close to the data, as demonstrated

We utilize a Difference-in-Difference approach to study the impact of HTE status on firms' innovation behaviors.[28] Table C.2 summarizes the results. With all other circumstances equal, an HTE firm tends to produce 14.0% more patents (column [1]), which we dub the "quantity effect". Meanwhile, its patents' quality drops by 24.3% (column [3]), which we dub the "quality effect" in relative terms. In the calibration exercise detailed in the following section, we use the magnitude of these two effects to discipline the values of two parameters: extra HTE subsidy, Δb_n, and the elasticity of innovation on entrepreneurial time, σ.

4.1.2 Parameters and Moments

In addition to the extension to HTE and non-HTE firms detailed in the previous section, we further include two extra policy parameters to the calibrated model: corporate tax rate u — hence firm static profit changing from πq to $(1-u)\pi q$ — and the R&D tax credit multiplier b_r, i.e., the total R&D cost changing from $R(x_d, x_m)$ to $(1 - b_r u)R(x_d, x_m)$.

The extended model has 23 parameters. We start with those that can be externally calibrated, directly inferred, or taken from the literature. We set the time discount rate ρ to match an annual interest rate of 6.2%.[29] For the inverse intertemporal substitution elasticity, we set $\nu = 4$ 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 20% 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 set the innovation elasticity parameter $\phi = 0.5$. We assume that the initial step-size of incremental innovations $\eta = \alpha\lambda$, and set the diminishing effect parameter $\alpha = 0.9$ following Akcigit and Kerr (2018). In the education sector, we set the death rate d so that an individual works for 35 years.

As for the subsidy-to-profit ratio, b_n, the theoretical model explains that it captures all explicit and implicit quantity-based subsidies a firm receives; however, in practice, we have to rely on observable subsidies to discipline it. We utilize the ratio of government subsidy

later in Table 4.4.

[28]The detailed explanation of the empirical model and strategies is provided in Appendix C.1.

[29]Values used in related literature vary from 4% (Storesletten et al., 2019) to 8% (Garriga et al., 2021) and ours lies between.

to profit[30] of public firms in four innovating sectors, *information technology*, *raw materials*, *industry*, and *communication services*, which amounted to 19.82% in 2012.[31] We therefore set $b_n = 19.82\%$. We set u and b_r to match a 25% corporate tax rate and a 150% tax credit multiplier in China. The values of all externally calibrated parameters and their sources are summarized in Table 4.1.

Table 4.1: Externally Calibrated Parameters

Para	Value	Equation	Meaning	Source
Aggregate Economy				
ρ	0.062	(1)	time discount rate	annual interest rate
ν	4	(1)	intertemporal elasticity of substitution	literature
ϵ	0.2	(2)	E.o.S. in final good production	profitability
L	1		total population	normalization
R&D Sector				
ϕ	0.5	(7)	innovation elasticity w.r.t. R&D	literature
α	0.929	(10)	diminishing effect	literature
η	$\alpha\lambda$	(10)	initial step-size of incremental inno.	assumption
Education Sector				
d	0.0286		death rate of the population	years of working
Government Policies				
b_n	19.82%	(13)	quantity-based subsidy	industry average
u	25%		corporate tax rate	documentations
b_r	150%		R&D tax credit multiplier	documentations

The remaining 12 parameters are internally calibrated to moments, which, unless stated otherwise, are calculated from our firm-level R&D input-output dataset in 2011-2013. As mentioned, we use the quantity and quality effects from HTE status to discipline the extra HTE subsidy, Δb_n, and the elasticity on entrepreneurial time, σ.

[30] More specifically, we use "net profit" from a firm's financial statement. Alternative choices are "operating profit" or "total profit", which are not so different in scale. For example, among public firms in 2012, the denominator-weighted average ratio of net to operating profit was around 82%, while that of the net to total profit was around 80%.

[31] The ratio for these four industries in the order listed in the main text is 29.13%, 41.07%, 12.64%, and 10.89%, respectively. The other sectors and subsidy-profit ratios (in brackets) are public affairs (15.82%), medicare (8.19%), real estate (4.86%), daily consumption goods (6.56%), energy (1.89%), and non-daily consumption goods (N/A).

The first set of remaining parameters regards innovation productivity. The model contains three productivity 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 spending to value-added,[32] of high- and low-type firms, which amount to 18.34% and 14.87% in the data. The ratio $z_{\mathrm{H}d}/z_{\mathrm{H}m}$ affects the share of radical innovations high-type firms choose to pursue. Therefore, we use the share of high-quality patents to discipline the value of $z_{\mathrm{H}d}$. We use the ASIE-Patent sample instead of 2008-2013 final analysis sample to incorporate the quality drop between pre- and post-2008 periods. Please refer to Appendix A.3 for detailed descriptions of data samples. The average share of high-quality patents was 13.50% in 2005-2008, dropping by 6.87 percentage points to 6.63% in 2011-2013.[33]

The second set of parameters regards skill intensities in innovation: γ_d and γ_m. Recall that low-type firms are capable of pursuing only low-quality patents, which embody incremental innovations. We use the observed skill intensity of low-type firms to discipline the skill intensity parameter in incremental innovations, γ_m. The skill intensity of high-type firms combines that in creating high- and low-quality patents. With the value of γ_m determined, we can discipline γ_d by targeting the observed skill intensity of high-type firms.

Based on equation (27), we set the step-size of radical innovations, λ, to match the annual TFP growth rate of 1.97% from 2008-2014, estimated by Bai and Zhang (2017). We use the skill premium to discipline the productivity in the education sector ξ, which determines the total supply of skilled labor and the equilibrium wage rates.[34] The fixed amount of unskilled labor employed in the education sector, ℓ^{edu}, is chosen to match the ratio of total higher education cost to GDP, 1.45%.[35]

[32] In the model, firms with the same innovation productivity choose the same level of R&D spending, but their value-added might differ due to idiosyncratic draws of product quality. "Average" means a within-type semi-aggregation which gives a "representative" value added for each type of firm.

[33] We also calculate this moment using the Innography sample, where the number is 6.25% in 2011-2013, similar to the ASIE-Patent sample.

[34] To obtain this value, 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.

[35] From Chinese Family Panel Studies, the share of household expenditure in income was 0.34% in 2012. The government expenditure on higher education to GDP ratio in the same year is 1.11%. The value we use is a summation of these two numbers.

Consistent with the model's settings, entry rate x_{E} is disciplined by the share of patents created by new entrants in the economy, 21.00%.[36] The cost coefficient for entrants to become high-type firms, χ, is set to match the percentage of high-type firms among all incumbents, 22.34% from 2011-2013.

In the end, we jointly calibrate all these 12 parameters to minimize the total sum of distance between model-generated and data moments

$$\sum_i \left| \frac{\text{model}(i) - \text{data}(i)}{\text{data}(i)} \right| \times 100\%.$$

Table 4.2 summarizes the internally calibrated parameters and their corresponding target moments.

Table 4.2: Internally Calibrated Parameters

Para	**Equation**	**Meaning**	**Target**
R&D Sector			
$z_{\mathrm{H}d}$	(7)	H-type firms' radical innov. prod.	share of radical innov.
$z_{\mathrm{H}m}$	(9)	H-type firms' incremental innov. prod.	H-type firms' R&D intensity
$z_{\mathrm{L}m}$	(9)	L-type firms' incremental innov. prod.	L-type firms' R&D intensity
γ_d	(7)	skill intensity in radical innov.	H-type firms' skill intensity
γ_m	(9)	skill intensity in incremental innov.	L-type firms' skill intensity
Other Sectors			
λ	(8)	step-size of radical innovations	TFP growth rate
ξ	(12)	education productivity	wage premium
ℓ^{edu}	(20)	unskilled labor in education	education cost to GDP ratio
x_{E}	(16)	entry rate	entrants' patent share
χ	(15)	cost of becoming H-type	fraction of H-type incumbents
Policy Variation			
σ	(7)	elasticity on entrepreneurial time	quality effect of HTE
Δb_n		quantity-based subsidy bonus for HTE	quantity effect of HTE

[36] Calibrated this way, the model implies an entry rate among innovating firms of around 15%, which is slightly higher than that among Chinese firms in the data. The exogenous entry rate plays a limited role in our quantitative results.

4.2 Model Fit

Targeted Moments. Table 4.3 presents the calibration results and model fit. The benchmark model well replicates the targeted moments. Among all 12 moments, only one has a relative distance exceeding 1 percentage point (p.p.). The total sum of data-model distance is 3.65 p.p. or an average distance of 0.30 p.p. per moment.

Table 4.3: Benchmark Calibration

Para	Value	Data Moment	Model Moment	Distance
$z_{\mathrm{H}d}$	1.0406	6.63%	6.62%	
$z_{\mathrm{H}m}$	1.0147	18.34%	18.19%	
$z_{\mathrm{L}m}$	0.9864	14.87%	14.68%	1.26 p.p.
γ_d	0.7808	31.30%	31.53%	
γ_m	0.4412	24.52%	24.51%	
λ	0.1306	1.97%	1.97%	
ξ	0.0337	2.43	2.43	
ℓ^{edu}	0.0094	1.45%	1.45%	
x_{E}	0.0633	21.00%	20.99%	
χ	0.2067	22.34%	22.32%	
σ	0.9729	-24.30%	-24.31%	
Δb_n	6.64%	14.00%	14.01%	

Note: Distance is calculated as the relative percentage deviation, and is provided only when exceeding 1 p.p.

Estimates of γ_d and γ_m imply that R&D activities pursuing radical innovations rely more heavily on skilled labor than incremental ones. The relatively significant difference between the two intensities, 0.7808 vs. 0.4412, is necessary to account for the observed 7% gap in the skill intensity between high- and low-type firms, as more than 85% of patents created by high-type firms are still of low-quality.

Our estimate of λ implies that a radical innovation improves the quality of a product by 13.06%. This number is close to that obtained in the literature. For example, Akcigit and Kerr (2018) estimated a step size of 11.20%, while Acemoglu et al. (2018) reported a step size of 13.20%. Furthermore, our model generates $\bar{\eta} = 0.0606$, that is, an incremental innovation on average improves the quality of a product by 6.06%.

The estimated value of σ indicates a large elasticity of innovation creation on entrepreneurial time. Discussions following Proposition 2 show that the quantity-quality trade-off strengthens as σ approaches its upper bound, $(1-\phi)/\phi$. Since we pick $\phi = 0.5$ from the literature, a value of $\sigma = 0.9729$ therefore suggests a strong trade-off on the micro-level, and an impact coefficient of 72.80 as specified in equation (24).

Non-targeted Moments. The first non-targeted moment we check is the aggregate creative destruction rate, which is challenging to measure from the data (Garcia-Macia et al., 2019). As we correspond 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. New patents in year t are those applied in year t and eventually granted. Patent stock refers to all active patents in year t.[37] For the 2011-2013 sample period, we estimate a patent-level creative destruction rate of 31.66%, which is close to the model counterpart $\delta = 30.16\%$.[38]

We then check the model's performance under the assumption of a random and permanent recognition of HTEs upon entry. Under the parameterization of $p^{\text{HTE}} = 55\%$, the model successfully generates a type distribution across incumbent firms close to the data, as demonstrated in Table 4.4. Specifically, our model predicts that 61.80% (58.40%) of high-type (low-type) incumbents are HTEs, while the data counterpart is 63.03% (59.48%). Furthermore, the HTE shares among incumbents are larger than the initial exogenous probability of recognition; HTEs are more active in innovating and expanding and are less likely to exit due to creative destruction. We consider the modeling of HTEs a success for capturing this important feature.

Table 4.4: Type Distribution of Incumbent Firms

High-Type	HTE	Model	Data
1	1	13.80%	14.08%
1	0	8.53%	8.26%
0	1	45.36%	46.19%
0	0	32.31%	31.47%

Note: This table reports the percentage of incumbent firms by their innovation productivity types and HTE status. The data column is calculated from sample period 2011-2013.

The model also predicts that firms with higher innovation intensity have a larger ex-

[37] Consistent with the patent distribution in Figure 4.1, here we use the Innography-ASIE matched sample of patents. Appendix C.3 provides a detailed description of variable construction.

[38] The number is higher than what is typically reported in the literature, especially those from the literature on firm dynamics; however, since we look at patent-level creative destruction and China experienced a patent surge in that period, we consider the high rate reasonable.

pected size. Under the calibrated parameter values, the size ratio between average high- and low-type firms, measured by employment, revenue, or profit, is 1.5285. Table 4.5 shows the relative size ratio from 2011-2013 firm-level data. Our calibration also captures the size difference well.

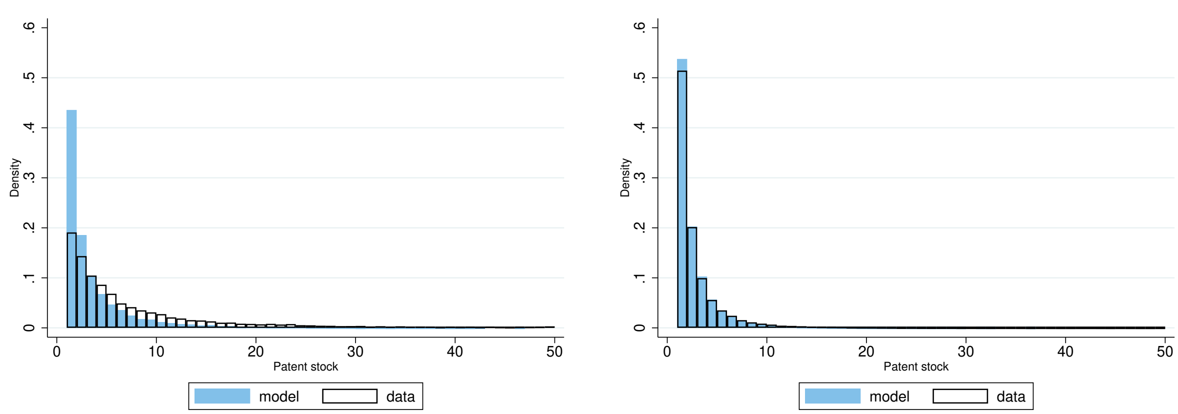
Table 4.5: Size Ratio between High- and Low-Type Firms in Data and Model

	Employment	Revenue	Profit
Data	1.2523	1.5044	1.6790
Model	1.5285	1.5285	1.5285

Note: This table reports the relative ratio for variables of interest, between average high- and low-type firms in 2011-2013 period, and we trim the bottom and the top 1 percent of the sample.

Lastly, we 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 4.1 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 4.1: 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 number among high- (panel (a)) and low-type (panel (b)) firms. Patent stock is calculated as the sum of all active patents within 2011-2013 period, and the distribution of patent stock is then estimated for the two sub-groups.

4.3 Effects of Quantity-Based Subsidies

We are now ready to evaluate the impact of quantity-based subsidies, the uniform subsidy b_n for all firms, and the extra subsidy Δb_n for HTEs. To that end, we compare the baseline outcome with a counterfactual economy in which all quantity-based subsidies are shut down, i.e., $b_n = \Delta b_n = 0$.

Table 4.6 presents the results. Compared to the counterfactual economy, quantity-based subsidies implemented in China generate an increase in innovation quantity from 0.2084 to 0.2383, equivalently by a relative 14.56%. Conversely, such subsidies reduce overall innovation quality, i.e., the share of radical innovations among total innovations, from 11.62% to 6.62%, or a decrease of 5 percentage points. In the data, compared to the pre-2008 period, we see a relative increase of 34.57% in patent quantity and a drop of 6.87 percentage points in patent quality in the post-2008 period.[39] Our calibration suggests that quantity-based subsidies account for 41.51% of the patent surge and 72.78% of the quality drop observed in the post-2008 period.

Table 4.6: 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}	$\frac{\Delta_{\text{Model}}}{\Delta_{\text{Data}}}$
$\delta - x_{\text{E}}$	innovation quantity	0.2383	0.2084	14.56%	34.57%	41.51%
δ_d / δ	innovation quality	6.62%	11.62%	-5.00%	-6.87%	72.78%

Note: Δ_{Model} represents change from the counterfactual to the benchmark economy. Δ_{Data} refers to change in the corresponding moments between the pre- and post-2008 period. Change in innovation quantity is in relative terms; while change in innovation quality, already measured in percentage terms, corresponds to the direct difference.

As detailed by equation (25), the drop in overall quality, induced by quantity-based subsidies, could come from the intensive margin as high-type firms face a quantity-quality trade-off or the extensive margin where firms endogenously choose types. Among the 5 percentage points decrease in patent quality shown in Table 4.6, we find that the intensive margin alone explains 142.70%, while that from the extensive margin is -20.62%. The reason is that quantity-based subsidies widen the value difference between high- and low-type firms, encouraging entrant firms to spend more on overhead investment

[39] The increase in quantity is estimated as a surge above the trend; details about the estimation can be found in Appendix D.1. The drop in quality is calculated from the ASIE-Patent sample between 2005-2008 and 2011-2013.

and become high-type firms. The economy now consists of more firms capable of pursuing radical innovations but creating mostly incremental ones. This situation corresponds to the scenario that Chinese firms and research institutes spend heavily on R&D infrastructures and laboratories but create very few high-quality patents, giving rise to another layer of resource waste.

By changing firms' innovation incentives, quantity-based subsidies further affect aggregate growth and welfare. In our exercise, the aggregate growth rate reduces from 2.30% in the counterfactual economy without subsidies to 1.97% in the baseline. Bai and Zhang (2017) report that Chinese TFP growth rate decreases by 1.91 percentage points, from 3.88% in 2001-2007 to 1.97% in 2008-2014. A drop of 0.33 percentage points in our model's TFP growth rate accounts for about 17% of the change between the data's pre- and post-2008 periods.

We follow equation (28) and decompose such effects into three channels, as shown in Table 4.7. Subsidies raise innovation quantity, leading to a 0.25% increase in aggregate growth rate; however, this positive quantity effect is overwhelmed by the negative quality effects. A pool of proportionately less radical innovations, which have a larger impact on productivity, reduces growth. This channel brings a 0.07% drop in aggregate growth. In addition, the average productivity enhancement from incremental innovations falls as quantity-based subsides induce more incremental R&D trials. This last channel brings a 0.51% drop in aggregate growth rate. Overall, the quality-crowding effect dominates, generating an overall negative effect on growth. As a result, eliminating quantity-based subsidies in the counterfactual economy creates a welfare gain of 9.39%.[40]

Table 4.7: Growth Decomposition

Δ_{Growth}	(i) quantity	(ii) quality-composition	(iii) quality-crowding	
-0.33	0.25	-0.07	-0.51	$(\times 10^{-2})$
	-75.24%	20.88%	150.53%	

Note: Growth differential is the growth rate difference between a counterfactual economy without quantity-based subsides and the baseline economy. We follow equation (28) and decompose the growth differential into three channels.

[40] As shown in equation (26), total welfare is jointly determined by initial consumption C_0 and the growth rate. In our exercise, the impact of C_0, with its small change, is secondary; the change in welfare comes predominantly from changes in the growth rate.

To summarize, the key channel for the negative growth and welfare effects of quantity-based subsidies is that although they promote overall innovations, such policies also worsen the quantity-quality trade-off facing innovating firms. The dominance of the quality effects, though, depends on our calibration of parameters, especially on σ, which governs the degree of quantity-quality trade-off. As a robustness check for our conclusion, we give a lower bound value of σ, above which the negative quality effects always dominate: $\underline{\sigma} = 0.9332$, or equivalently, an impact coefficient of 28.94.[41] More details can be found in Appendix D.2.

4.4 Social Planner's Allocation

This section analyzes the scenario of a constrained planner's allocation. 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 our economy contains an education sector with endogenous creation of skilled labor, the planner only needs to choose the threshold talent above which the young shall obtain education, θ^*, to maximize social welfare. Once θ^* is chosen, skilled and unskilled labor supplies are determined.[42]

We detail how to solve the social planner's problem in Appendix D.3. The aggregate welfare is a hump-shaped function of the education threshold, θ^*. An increase in skill supply initially raises welfare as it promotes innovation and growth but reduces welfare after passing a specific range, as the negative effect from a shrinking unskilled workforce and lower initial consumption level eventually dominates. Table 4.8 compares the benchmark economy and the social planner's allocation.

The results suggest a rather large room for welfare improvement. The socially optimal level of skilled labor supply is more than twice that of its counterpart in the market equilibrium, which dramatically reduces the skill premium by about 37%. As more skilled labor promotes innovation, the aggregate growth rate increases from 1.97% to 4.23%, and the aggregate welfare improves by 17.73%.

The aggregate welfare gain is affected by the parameter governing the elasticity of intertemporal substitution, ν. In the benchmark calibration, we set $\nu = 4$, which delivers a

[41] The benchmark calibration of σ amounts to 0.9729, with a much higher impact coefficient of 72.80. As shown in equation (24), the impact coefficient varies drastically when σ approaches the upper bound.

[42] See Appendix B.1 for details of how skill supply is determined with a given value of θ^*.

Table 4.8: Comparison between Market Equilibrium and Social Planner's Allocation

Variable	Meaning	Benchmark	Planner
θ^*	education threshold	5.12	3.14
h^{supply}	skill supply	3.71%	9.84%
w^h/w^ℓ	skill premium	2.43	1.52
g	TFP growth rate	1.97%	4.23%
C_0	initial cons. level	100%	96.91%
U	social welfare	100%	117.73%

Note: The benchmark market equilibrium level of C_0 and U are normalized to 100%, respectively. h^{supply} represents both the fraction and the level, since we normalize the total population of workers $L = 1$.

relatively small welfare elasticity concerning the growth rate. In Appendix D.3, we show welfare gains under different values of ν.

4.5 Skill Subsidies

We analyzed the welfare gains from allocations chosen by a constrained planner. A natural follow-up question is whether policymakers can find easy-to-implement subsidies to recover the planner's allocation and improve social welfare. One policy is to design more "selective" innovation subsidies, for example, subsidies to input and output of radical innovations only. Chen et al. (2021) document that many Chinese firms fake reporting R&D expenditure to qualify for government subsidies. In reality, it is even harder to require firms to truthfully report the kind of innovations they are pursuing.

Here we propose a second, but equally effective policy: *skill subsidies*, or *subsidizing the human capital accumulation*. Skill subsidies are biased toward high-quality innovations since they are more skill-intensive. In particular, we incorporate an education subsidy into the model with a policy parameter $b_e \in [0, 100\%]$.[43] Recall that the total education cost is $\left(w^h h^{\text{teacher}} + w^\ell \ell_{\text{edu}}\right)$. With education subsidy, a b_e portion of total education cost is covered by the government. A young individual's threshold talent for attaining education

[43] The model contains two equivalent forms of skill subsidies: *education subsidy* which covers part of education cost, and *skilled labor subsidy*, which covers part of the firms' wage cost of hiring skilled labor.

becomes

$$\theta^* \equiv \max \left\{ \frac{1-b_e}{\xi} \left[1 - e^{-(\rho+d)} \right] \left(e^{-(\rho+d)} - \frac{w^\ell}{w^h} \right)^{-1}, 1 \right\}. \tag{29}$$

Consequently, the larger the education subsidy b_e is, the lower the threshold θ^* will be, and more young workers will obtain an education, raising the skilled labor supply.

Efficiency of Skill Subsides. In Section 4.4, we show that compared to market equilibrium, the constrained efficient allocation features a larger skill supply and lower skill premium. The issue is, with such a low skill premium, not that many individuals would willingly choose to invest in education and become skilled. At a skill premium of 1.52, the implied level from the planner's allocation, individually rational choices on education yields a significantly higher talent threshold of $\theta^* = 10.08$, and correspondingly, a much lower skill supply accounting for only 0.96% of the population.

Not surprisingly, education subsidies can induce people to obtain education and become skilled labor even with a low skill premium.[44] Under our calibrated parameter values, an education subsidy that recovers the planner's allocation is $b_e = 68.83\%$. That is, if the government compensates for 68.83% of the education cost, at a low skill premium of 1.52, any individual with talent above the socially optimal threshold, $\theta^*_{\text{SP}} = 3.14$, can optimally choose to obtain an education.

An education subsidy rate $b_e = 68.83\%$ implies that the ratio of government education expenditure to GDP is 1.95%, which is significant considering that in the benchmark economy, the total size of the education sector was only 1.45% of the GDP. Therefore, we regard $b_e = 68.83\%$ not as a policy recommendation to be implemented immediately but instead as a measurement of room for improvement.[45]

Skill Subsidy as a Quality-Biased Policy. Here, we focus on illustrating why education subsidy achieves a better outcome than quantity-based subsidy. To that end, we start from the benchmark economy with quantity-based subsidy $b_n = 19.82\%$ and education subsidy $b_e = 0$ and compare it with two counterfactual situations with 5 percentage points strengthening each subsidy. Table 4.9 summarizes the results.

[44] From equation (29), it is straightforward to see that if we put θ^*_{SP} on the left-hand side and the skill premium on the right, there always exists a b_e such that the equality holds.

[45] Note that in the model, improving the productivity of the education sector, ξ, which is arguably more challenging to implement in the short run, has the same effect as education subsidies.

Table 4.9: Comparison between Quantity-Based Innovation Subsidies and the Education Subsidy

Variable	Meaning	Benchmark	b_n+5%	b_e+5%
$R(x)/Vadd$	average R&D intensity	15.75%	16.71%	15.85%
x_{Hd}/x_H	% of radical inno. by H-type firms	25.21%	21.62%	28.95%
w^h/w^ℓ	skill premium	2.4318	2.4745	2.3577
p^*	fraction of H-type entrants	18.94%	19.40%	19.30%
g	TFP growth rate	1.97%	1.87%	2.17%
C_0	initial consumption level	100%	99.40%	99.87%
U	social welfare	100%	97.41%	102.58%

Note: The benchmark market equilibrium level of C_0 and U are normalized to 100%, respectively. x_H is calculated by aggregating x_H^{HTE} and $x_H^{non\text{-}HTE}$, same for x_{Hd} and x_L.

As mentioned in Section 4.3, an increase in quantity-based innovation subsidies, though promoting aggregate R&D (row [1] in Table 4.9), worsens firms' choices in their quantity-quality trade-off (row [2]); a rise in the skill premium (row [3]) amplifies this negative effect. Consequently, both the aggregate growth rate (row [5]) and the welfare (row [7]) are reduced.[46] Different from quantity-based innovation subsidies, an education subsidy is quality-biased as radical innovations are more skill-intensive than non-radical ones. An increase in education subsidy raises the skill supply. It reduces skill premium, encouraging firms to pursue proportionately more radical innovations, which increases both growth and welfare.[47]

5 Conclusion

This paper studies the growth and welfare implications of quantity-based innovation subsidies widely adopted in China. To that end, we first construct an innovation input-output dataset for the Chinese industrial firms, which serves as a valuable tool-set to

[46] An increase in quantity-based subsidy (as well as education subsidy) raises the fraction of high-type entrants (row [4]) and reduces the initial consumption level (row [6]). We provide corresponding results with strengthening R&D tax credit b_r in Appendix D.4.

[47] In the model, skilled labor is employed only in R&D. There is a concern that, in reality, an increase in the skilled labor supply may not be absorbed totally by the R&D sector, but also to production activities, which weakens the growth-stimulating effect of skill subsidies. Thus, we estimate the demand elasticity for skilled labor of innovating vs. non-innovating firms and find that elasticity of the former is three times higher, confirming that a reduction in skill premium benefits R&D activities more. Appendix D.5 provides details about the estimation.

study micro-level innovation decisions in the Chinese economy. We empirically document a rapid quantity surge and a steady quality decline of Chinese patents since the mid-2000s and highlight the importance of skilled labor in R&D activities pursuing high-quality patents.

Motivated by these facts, we build a Schumpeterian growth model featuring heterogeneous innovations: radical vs. incremental. By introducing a scarce R&D resource, the model generates a quantity-quality trade-off between radical and incremental inventions facing innovative firms. Moreover, the model allows us to decompose the effects of innovation subsidies into quantity and quality channels. We calibrate the model to our firm-level innovation dataset. The model-based quantitative analysis shows that quantity-based subsidies reduce the aggregate welfare by about 9%, as the subsidies induce negative quality channel effects which are dominant. Within the framework, we propose and evaluate a quality-biased and welfare improving policy — subsidizing the human capital accumulation. This latter policy raises the supply of skilled labor, benefits the more skill-intensive radical innovations, and improves both the quantity and quality of innovations.

We necessarily abstract from other essential features to focus on the quantity-quality trade-off, which we view as important in studying Chinese innovations given its patent surge. For example, we do not model the imitation to innovation transition in China. Peters and Zilibotti (2021) develops a model that incorporates a world technology frontier into a Schumpeterian growth framework with heterogeneous firms to study innovation and industrial policies in developing countries. Extending our model to allow for distance to the world's frontier and examining its interaction with innovation quality would be an interesting avenue for future research.

While our model speaks to the importance of subsidizing the human capital accumulation, the way it is modeled is simplified to keep the framework tractable. A more detailed investigation into the impact of skill accumulation is a natural extension. For example, patent subsidies work immediately, while building up a skill pool takes generations of time. Investigating this dimension of heterogeneity would provide a more accurate evaluation of policies. From the perspective of policy implementation, various forms of human capital subsidies exist in the real world. In addition to subsidizing education costs, which we emphasized in the model, attracting overseas-trained talents to work in universities and the industrial sector might be another important channel for China's technology catch-up. We leave these studies 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*, 2020, *forthcoming.*

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

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica: Journal of the Econometric Society*, 1992, pp. 323–351.

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.

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.

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.

Chen, Hong, Hanbing Fan, Takeo Hoshi, and Dezhuang Hu, "Do Innovation Subsidies Make Chinese Firms More Innovative? Evidence from the China Employer Employee Survey," Technical Report, National Bureau of Economic Research 2019.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu, "Notching R&D Investment with Corporate Income Tax Cuts in China," Technical Report, National Bureau of Economic Research 2021.

Chen, Zhiyuan and Jie Zhang, "Types of Patents and Driving Forces behind the Patent Growth in China," *Economic Modelling*, 2019, *80*, 294–302.

Dang, Jianwei and Kazuyuki Motohashi, "Patent Statistics: A Good Indicator for Innovation in China? Patent Subsidy Program Impacts on Patent Quality," *China Economic Review*, 2015, *35*, 137–155.

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.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J. Klenow, "How Destructive is Innovation?," *Econometrica*, 2019, *87* (5), 1507–1541.

Garriga, Carlos, Aaron Hedlund, Yang Tang, and Ping Wang, "Rural-Urban Migration and House Prices in China," *NBER Working Paper*, 2021.

Grossman, Gene M and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *The review of economic studies*, 1991, *58* (1), 43–61.

He, Zi-Lin, Tony W Tong, Yuchen Zhang, and Wenlong He, "Construction of a Database Linking SIPO Patents to Firms in China's Annual Survey of Industrial Enterprises 1998-2009," Technical Report, Tilburg University working paper 2016.

_ , _ , _ , and _ , "Constructing a Chinese Patent Database of Listed Firms in China: Descriptions, Lessons, and Insights," *Journal of Economics & Management Strategy*, 2018, *27* (3), 579–606.

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.

Jia, Junxue and Guangrong Ma, "Do R&D Tax Incentives Work? Firm-Level Evidence from China," *China Economic Review*, 2017, *46*, 50–66.

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?," Technical Report, National Bureau of Economic Research 2020.

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.

Peters, Michael and Fabrizio Zilibotti, "Creative Destruction, Distance to Frontier, and Economic Development," *NBER Working Paper*, 2021.

Storesletten, Kjetil, Bo Zhao, and Fabrizio Zilibotti, "Business Cycle during Structural Change: Arthur Lewis' Theory from a Neoclassical Perspective," *NBER Working Paper*, 2019.

Sun, Zhen, Zhen Lei, Brian D Wright, Mark Cohen, and Taoxiong Liu, "Government Targets, End-of-Year Patenting Rush and Innovative Performance in China," *Nature Biotechnology*, 2021, *39*, 1068–1075.

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, Shangjin, Jianhuan Xu, Ge Yin, and Xiaobo Zhang, "Bureaucratic Bean Counting and Patent Subsidies: A Welfare Evaluation of China's Pro-Innovation Program," *Working Paper*, 2021.

Appendix of "Haste Makes Waste? Quantity-Based Subsidies under Heterogeneous Innovations"

A Data Source and Construction

A.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), contain balance sheet information for all Chinese industrial firms with sales above 5 million RMB before 2011 and 20 million RMB since 2011 (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 have been applied and eventually granted in China. For instance, if a patent is filed in China in year t and eventually granted, it consists of our sample of newly created Chinese patents in year t.

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 2008-2013 period. There are in total 394,381 observations within the seven-year time period, covering over approximately 120,000 unique firms.

A.2 Variable Construction

High-Quality Patents. In the benchmark, we define a patent as a high-quality patent (HQ patent) if it has ever been cited by a US patent, which can be observed from the Innography Database. As long as one of its forward citations is from the United States, then this patent is classified as a high-quality patent. As the robustness checks, we define a patent as "high-quality" if its patent family size is greater than one (i.e., an invention applies for patents both in China and at least in one foreign country.).

Skill Composition. We define the employment engaging in scientific activities (*keji huodong renyuan*) as R&D personnel, and among all R&D personnel we further categorize those with medium or high professional titles (*zhonggaoji zhicheng*) as skilled labor. Skill intensity is then defined as the ratio of skilled labor and total R&D personnel. However, for the regression analysis using 2008-2013 sample, as information of workers with medium or high professional titles is not available in 2008 and 2009, instead, we use workers with bachelor and above degree to proxy for the skilled. And we define the skilled intensity in these two years as the ratio of the number of workers engaging in scientific activities with a bachelor's degree or above and R&D personnel.

Entrants, Incumbents, and Innovating Firms. In the full sample period 1998-2013, the first time that one firm starts to apply (and eventually own) at least one patent is identified as its entry year into innovating, and the firm itself is defined as a new entrant in that given survey year. Regardless of whether to own patents or not in subsequent years, a firm is identified as an innovating firm starting from its entry year and as an incumbent starting from the second year since it entered the innovation market. For example, if a firm had at least one patent in 2001 but none in other years, it will be characterized as non-innovating in 1998-2000, entrant in 2001, innovating in 2001-2013, and incumbent in 2002-2013.

High-Tech Enterprises. Firm Innovation Activity Database reports tax exemption amount from the Recognition of High-Tech Enterprises. We define a firm to be a HTE firm if it has a positive value of tax exemption amount, otherwise it is classified as non-HT firm.

A.3 Sample Construction

To construct the micro-firm-level samples, we need to employ data from various sources and merge different data sources using firms' Chinese names. The sample construction process consists of the following three major steps.

Step 1: Construct 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 consists of 4,037,866 firm-level observations.

Step 2: Attach Patent Information to ASIE Sample using Patent Applicant Information — "ASIE-Patent sample". Firm-level ASIE data and patent-level Innography Database are merged by using information on institutional applicants of patents. When calculating changes between pre-2008 and post-2008 periods, we utilize this ASIE-Patent sample.

Step 3: Attach Firm-Level Innovation Activities Data to ASIE-Patent Sample — "final analysis sample". Last, we merge firm-level innovation activities data with ASIE-Patent sample using firms' names. Following He et al. (2016) and He et al. (2018), we prepare a list of "clean" firm names. In particular, we trim all symbols and punctuation marks, convert all full-width letters and numbers into half-width ones, convert Chinese numbers into Arabic numbers, remove various designators, and remove location levels within the names. Then we can use firms' cleaned names in Chinese to merge Firm Innovation Activity Database with the ASIE-Patent sample to construct our final analysis sample.

The following Table A.1 shows the sample size during the sample construction process.

Table A.1: Sample Construction

	Raw			Matched	
Year	ASIE	Innography	Innovation	ASIE-Patent	Final Sample
2005	249976	108600	—	2149	—
2006	260188	122471	—	2813	—
2007	287189	133902	—	3735	—
2008	371887	154444	30267	6286	3760
2009	341087	176648	37369	7775	4859
2010	415514	204728	42139	10692	7049
2011	280405	250180	53762	9652	9652
2012	301335	306421	67053	12522	12522
2013	321157	372926	76167	14873	14873
Data Type	firm-level	patent-level	firm-level	firm-level	firm-level

Note: This table reports the number of observations of various data sources during sample construction process.

A.4 Additional Figures and Tables

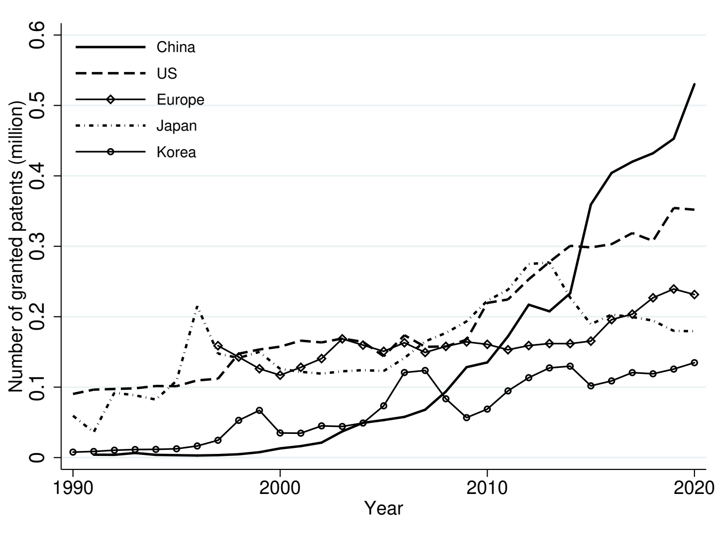


Figure A.1: Number of Granted Patents in China and Advanced Economies

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

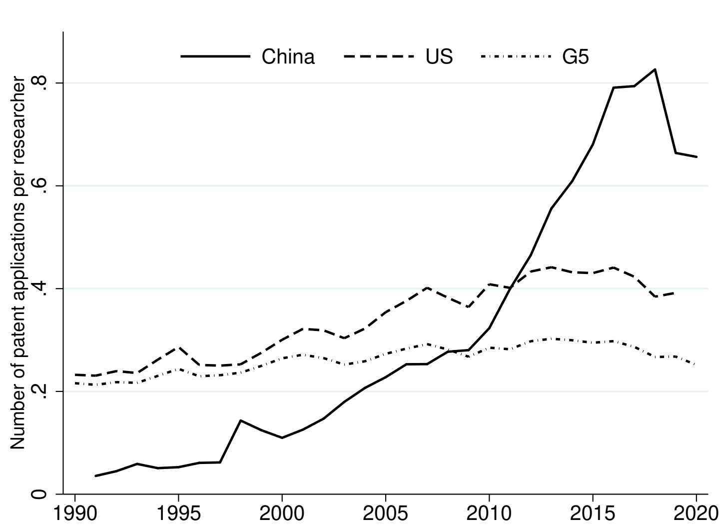
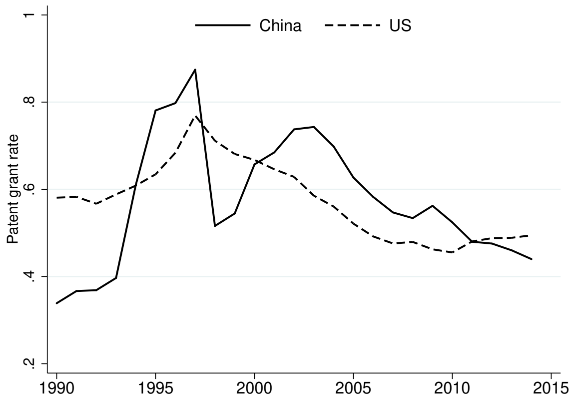


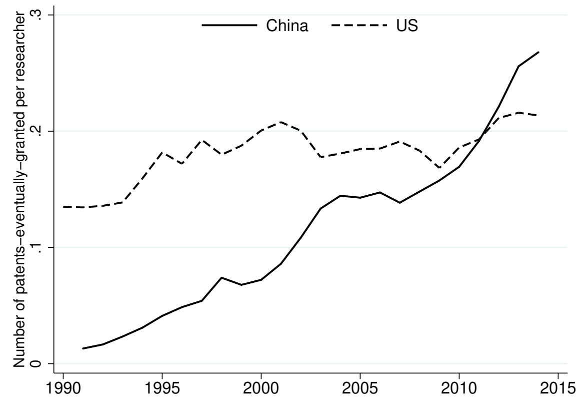
Figure A.2: Number of Patent Applications per Researcher in China, the US and G5

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 include 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 is 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.

Table A.2: Summary Statistics of the Final Analysis Sample

	2008	2009	2010	2011	2012	2013
Firm Age	14.86	15.15	15.23	14.89	14.90	14.73
	(14.78)	(14.03)	(13.40)	(12.77)	(12.39)	(11.79)
Foreign-Owned Enterprises (%)	25.74	26.61	28.06	24.69	24.29	22.46
	(43.73)	(44.20)	(44.93)	(43.12)	(42.89)	(41.74)
Private-Owned Enterprises (%)	68.62	67.48	67.40	70.47	71.40	74.91
	(46.41)	(46.85)	(46.88)	(45.62)	(45.19)	(43.36)
	full sample					
Firm-Level HQ Patent Share (%)	0.08	0.07	0.07	0.05	0.03	0.02
	(0.23)	(0.21)	(0.20)	(0.17)	(0.13)	(0.12)
	high-type sample					
Firm-Level HQ Patent Share (%)	0.26	0.25	0.24	0.19	0.15	0.14
	(0.34)	(0.33)	(0.32)	(0.29)	(0.26)	(0.25)
Number of Observations	3760	4859	7049	9652	12522	14873

Note: This table reports the summary statistics of the final analysis sample, merged from ASIE, Innography Patent Database, and Firm Innovation Activity Database. Means and standard deviations in parentheses.

Table A.3: Skill Intensity and Patent Quality

	high-type sample	high-type sample	high-type sample
	(1)	(2)	(3)
	HQ patent share	HQ patent share	HQ patent share
high-skill researcher share	0.164**	0.159**	0.159**
	(2.25)	(2.23)	(2.24)
R^2	0.283	0.305	0.324
Firm characteristics controlled	Yes	Yes	Yes
Industry F.E.	Yes	Yes	Yes
Year F.E.	No	Yes	Yes
Location, Ownership F.E.	No	No	Yes
Observations	3828	3828	3828

Note: This table shows results regressing the firm-level high-quality patent share on high-skill researcher share, controlling for firm characteristics with robust standard errors clustered at the firm-level. t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Model Derivations and Proofs

B.1 Education

Young people choose to invest in education and become skilled if and only if

$$\frac{e^{-(\rho+d)}}{\rho+d}w^h - \frac{1-e^{-(\rho+d)}}{\rho+d}\frac{w^h}{\theta\xi} \geq \frac{w^\ell}{\rho+d}.$$

A young worker chooses to invest in education if and only if her type is above a certain threshold

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

We can further derive the mass of the four types of workers 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 Product Line Distribution

Start with the step-size distribution of an incremental innovation. 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 to 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 product line 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 the j-type, 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 being 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 $N \geq 2$ many types of firms in the economy.[48] And define $\mathrm{Line}_j \equiv \sum_n \mu_{j,n} \times n$, that is, the total number of product

[48] $N = 2$ in our baseline model. Later on in the quantified model, we extend to $N = 4$. This is why we decide to prove a more general version here.

lines held by j-type firms.

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

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

We plug in the definition of δ, it becomes

$$\text{Line}_j = \frac{p_j x_{\text{E}}}{\sum\limits_i \text{Line}_i x_i + x_{\text{E}} - x_j}, \tag{eqn-[j]}$$

together with the requirement of

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

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

It seems that we need an extra free variable to such that it is a system of $N+1$ unknowns and $N+1$ equations. However, we are going to prove that, for any given combinations of $x_{\text{E}} > 0$, $\{p_j\}_{j=1}^N \in (0,1)$ and $\{x_j\}_{j=1}^N > 0$, there always exists a $\{\text{Line}_j\}_{j=1}^N$ such that the above $N+1$ equations hold.

The proof is to show that, when eqn-[1] to eqn-[N-1] hold and eqn-[x] is satisfied, the last equation, eqn-[N], shall hold automatically.

eqn-[j] indicates that

$$(p_j - \text{Line}_j)x_{\text{E}} = \text{Line}_j \left(\sum_i \text{Line}_i x_i - x_j \right). \tag{eqn-[j']}$$

Sum them up from j = 1 to N-1, and use the fact that $p_N = 1 - \sum\limits_{j=1}^{N-1} p_j$, $\text{Line}_N = 1 - \sum\limits_{j=1}^{N-1} \text{Line}_j$, we have

$$(\text{Line}_N - p_N)x_{\text{E}} = (1 - \text{Line}_N) \left(\sum_i \text{Line}_i x_i \right) - \sum_{j=1}^{N-1} \text{Line}_j x_j.$$

For the R.H.S., let's rearrange terms based on x_j, and get

$$(\text{Line}_N - p_N)x_{\text{E}} = (1 - \text{Line}_N)\text{Line}_N x_N - \text{Line}_N \sum_{j=1}^{N-1} \text{Line}_j x_j,$$

which is exactly the same as eqn-[N]

$$(p_N - \text{Line}_N)x_{\text{E}} = \text{Line}_N \left(\sum_{j=1}^{N} \text{Line}_j x_j - x_N \right).$$

B.4 Value Functions and Proof of Proposition 2

Without loss of generality, we focus on the value function of high-type firms. Guess that the value function takes the following form

$$V(Q, \bar{q}) = \sum_i A q_i + nB\bar{q}.$$

Substituting this conjectured form into the Bellman equation, we have

$$\begin{aligned} r\left(\sum_i A q_i + nB\bar{q}\right) - gnB\bar{q} = \max_{x_d, x_m} \; & \sum_i \left[\pi q_i - \delta(A q_i + B\bar{q})\right] + n x_d \left[A(1+\lambda) + B\right]\bar{q} \\ & + n x_m \left[A(1+\bar{\eta}) + B\right]\bar{q} - nR(x_d, x_m) + n b_n \pi \bar{q}. \end{aligned}$$

It follows that A and B satisfy the following conditions

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

and

$$(\rho + \delta)B = x_d\left[A(1+\lambda) + B\right] + x_m\left[A(1+\bar{\eta}) + B\right] - R(x_d, x_m) + b_n \pi.$$

One can see that B is increasing in b_n.

With the value function's form, equation (23) in Proposition 2 follows immediately from the first-order conditions of x_d and x_m.

C Calibration

C.1 InnoCom Program and HTE

InnoCom Program. China has initiated many subsidy programs aim to promote firm innovations around 2008. A critical such tool is the recognition of HTEs under the InnoCom Program.[49] From 2008-2014, HTEs account for 8.2% of total industrial employment and for 16.2% of total industrial value added in China. From 2013-2019, HTEs account for 56.9% of total R&D personnel in China.[50] In addition to a corporate tax cut from 25% to 15%, qualified HTEs can enjoy various types of research and development subsidies such as research grants and patent subsidies. At the same time, qualified HTEs are required to file reports on innovation behaviors, for example, share of R&D personnel and number of newly granted patents, with the local government every year to maintain the HTE status.

In practice, we label a firm as HTE if it enjoys a positive HTE tax exemption. As a consequence, there are four types of firms in the data: high-type & HTE; high-type & non-HTE; low-type & HTE; low-type & non-HTE. Table C.1 shows the distribution across types for entrant firms from 2011-2013. During the sample period, 53.40% high-type entrant firms are HTEs, while 55.60% low-type ones are HTEs.

Table C.1: Type Distribution of Entrant Firms

High-Type	HTE	Number	Percentage
1	1	550	53.40%
1	0	480	—
0	1	7473	55.60%
0	0	5967	—

Note: This table shows the number of entrant firms by their innovation productivity types and HTE status in the sample period 2011-2013.

Relationship between HTE and Firm Innovation. We utilize a Difference-in-Difference (DID) approach to study the impact of HTE recognition on firm innovations. We define a

[49] Among the qualifications to become a 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.

[50] This numbers are from China Torch Statistical Yearbooks.

post dummy equals to 1 for a HTE if the year of the observation lies on or after the year that the firm obtains the HTE title for the first time, and 0 otherwise. For a Non-HTE firm, we define post always equal to zero. A firm could have a US-cited patent after it has become HTE, we thus define a HTE firm as high-type only if it has at least one US citation before the first HTE recognition year in the DID specification. To do a meaningful before-after comparison, we keep only HTEs with at least one prior year and at least one post year (including the recognition year).

We then regress (log) patent quantity and (log) high-quality patent share on the HT dummy, the interaction term between HT dummy and post, which is the key variable of our interest, controlling for firm age, (log) employment, (log) revenue, (log) assets as well as year, location (province), industry, and ownership types fixed effects. Robust standard errors are clustered at the firm-level. It is notable that the post dummy is absorbed by our year fixed effects. Table C.2 shows the results. We find that after being successfully certified as HTEs, the number of Chinese invention patents that firms produce expands by 14% (column [1]) while the high-quality patent share shrinks by 24.3% (column [3]), both in relative terms.

Table C.2: HTE and Firm Innovation: DID analysis

	full sample	high-type firms sample	
	(1) patent quantity	(2) patent quantity	(3) HQ patent share
HT dummy	0.164***	0.146***	-0.114**
	(8.43)	(2.76)	(-2.44)
HT dummy × post	**0.140*****	0.117*	**-0.243*****
	(5.84)	(1.91)	(-3.60)
R^2	0.202	0.351	0.350
firm characteristics controlled			
Observations	30184	5731	3038

Note: This table shows results regressing the dummy variables for High-Tech firms and its interaction with the dummy variable denoting the before- and after- HTE recognition on firms' innovation activities such as (log) total number of patents for the full sample (column [1]) and (log) high-quality patent share for the selected high-type-firm sample (column [2]-[3]), controlling for year, location, industry, and ownership types fixed effects with standard errors clustered at the firm-level. Other firm characteristics are also controlled. t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

As pointed out by Dang and Motohashi (2015) and Li (2012), innovation subsidy programs usually levy a positive impact on patent quantity. Hence, it is not surprising that InnoCom Program has a positive influence on the quantity margin. It is mandatory for HTEs to report their innovation activities every year, which are also measured by the quantity of patents. This gives HTEs extra incentives to keep producing new patents every year (Sun et al., 2021). Besides, in order to maintain the HTE status, one pre-requisite is to employ at least 10% workers engaging in R&D activities. Last but not least, the status of HTE has to be renewed every three years, for which the number of patents is an important metric.[51] All of these factors motivate HTEs to increase the quantity of innovations. However, since quality is not targeted directly, and difficult to measure promptly, firms' incentive for pursuing high-quality innovations is weak, as reflected by the negative quality effect in Table C.2.

C.2 Alternative Definitions of High-Quality Patents

The quality criteria of being cited by US patents might be affected by other factors, such as China entering WTO or changes in policies regulating foreign citations. In this section, we perform a robustness check by defining high-quality patents based on the family size of a patent. In particular,we use European Patent Office's (EPO) DOCDB family ID to identify patent family, which we extract from EPO's PATSAT database. A Chinese patent is regarded as high-quality if its family size is greater than one. Essentially, A family size greater than one means the same invention applies for patent not only in China but also at least in another foreign country. Patent family size is widely used by innovation scholars as an indicator for patent quality.

Table C.3 presents the regression result, which shows that the positive quantity effect and negative quality effect of HTE recognition still hold under this alternative definition of high-quality patents.

C.3 Active Patent Stock

We construct the stock of active patents using the information of patents' forward citations for the Innography-ASIE matched sample. More specifically, we define the lifespan of a patent from its application year to the last year it has received a forward citation. The idea here is that new patents might obsolete some old ones, causing a creative destruction

[51] These requirements are listed in the Administrative Measures for Determination of High and New Tech Enterprises (*Gaoxin Jishu Qiye Rending Guangli Banfa*) 2008 version and 2016 version.

Table C.3: HTE and Firm Innovation (robustness check): DID analysis

	full sample	high-type sample	
	(1) patent quantity	(2) patent quantity	(3) HQ patent share
HT dummy	0.164***	0.130*	-0.173**
	(8.43)	(1.94)	(-2.36)
HT dummy × post	**0.140*****	0.0572	**-0.144***
	(5.84)	(0.85)	(-1.68)
R^2	0.202	0.391	0.368
firm characteristics controlled			
Observations	30184	3970	1996

Note: This table shows results regressing the dummy variables for High-Tech firms and its interaction with the dummy variable denoting the before- and after- HTE recognition on firms' innovation activities such as total number of patents for the full sample (column [1]) and high-quality patent share for the selected high-type-firm sample (column [2]-[3]), controlling for year, location, industry, and ownership types fixed effects with standard errors clustered at the firm-level. Other firm characteristics are also controlled. t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in the society's knowledge pool. As a simple but intuitive way, we regard an old patent "inactive" or "dead", when it no longer contributes to the society's knowledge creation.

To find the latest forward citation, we first list all forward citations for a given patent, rank them using their application years, and choose the latest one. In the end, for a granted patent which was applied in year t_0, if the application year associated with its latest forward citation is t, then this patent is regarded as "active" for the whole t_0 to t period. If a patent doesn't receive any forward citation, we assume that it is active only in its application year (alive for one period only) and becomes inactive in the subsequent periods. We then construct the patent-level creative destruction rate as the ratio of newly granted patents and stock of active patents in a given year, which yields the average rate of 31.66% in 2011-2013 period.

Figure 4.1. After constructing the stock of active patent, we further match the patent information back to the firm-level ASIE sample, and divide the sample into high-type firms and low-type firms. High-type firms are those with at least one high-quality patent, i.e., a patent that has been cited by US patents. Then, we use histogram to illustrate the density distribution of live patent stock for the two sub-samples in 2011-2013 sample period. For the illustration purpose, we restrict to the sample with patents less than 50.

D Quantitative Analysis

D.1 Estimation on the Magnitude of Patent Surge

It is clear from Figure 2.2 (and Figure A.3), the number of patents per researcher increases at a faster rate in the post-2008 period than the pre-2008 period. To estimate the magnitude of patent surge above its natural trend in the post-2008 period, we first fit the pre-2008 data of patents per researcher with a linear trend. Using the pre-2008 trend in patents per researcher and the number of researchers in years after 2008, we then obtain a predicted series of patent quantity, if patents shall grow at the same rate after 2008 as in the pre-2008 period. Then by comparing the actual number of patents in a post-2008 year with the predicted value, we obtain the estimation on the magnitude of patent surge above trend. There is no information on number of researchers among industrial firms before 2008. We assume that the share of researchers in industrial firms in total researchers is constant. Under this assumption, we first divide number of ASIE patents by national researchers, and then use this trend to predict subsidy-induced patents number in the post-2008 period. Table D.1 summarizes the estimation results, we use the number from 2011-2013 period, under a 2002-2008 trend among ASIE patents for Table 4.6.

Table D.1: Magnitude of Patent Surge

Trend	2010-2013	2011-2013
1998-2008	39.20%	45.62%
2002-2008	28.94%	**34.57%**

D.2 A Lower Bound Value of σ

In this section, we provide a lower bound value of σ, above which the negative quality effects always dominate the positive quantity effect, of quantity-based subsidies. In particular, we find $\underline{\sigma} = 0.9332$, with other parameters kept at their benchmark values. That is, as long as $\sigma \in (0.9332, 1)$, or equivalently, a range of $(28.94, \infty)$ for the impact coefficient specified in equation (24), quantity-based subsidies shall induce a negative growth effect in net.[52]

[52]The net effect on welfare differs from that on growth, but the difference is quantitatively negligible.

The following table replicates Table 4.6. Consistent with our discussions following Proposition 2, with a smaller σ, though the quantity effect stays nearly the same (row [1]), the subsidies now have much weaker quality effects (row [2]), and the net effect on TFP growth rate turns out to be 0 by design (row [3]).

Table D.2: Moments With and Without Quantity-Based Subsidies

Variable	**Meaning**	$\underline{\sigma}$	**C.F.**	$\Delta_{\textbf{Model}}$	$\Delta_{\textbf{Data}}$	$\frac{\Delta_{\textbf{Model}}}{\Delta_{\textbf{Data}}}$
$\delta - x_{\mathrm{E}}$	innovation quantity	0.2396	0.2102	13.99%	34.57%	40.47%
δ_d/δ	innovation quality	10.71%	13.97%	-3.26%	-6.87%	47.45%
g	TFP growth rate	2.48%	2.48%	0.00%	—	—

Note: Δ_{Model} represents change from the counterfactual to the original economy. Δ_{Data} refers to change in the corresponding moments between the pre- and post-2008 period. Change in innovation quantity is in relative terms; while change in innovation quality, already measured in percentage terms, corresponds to the direct difference.

D.3 Social Optimum under Different Values of ν

We've solved the planner's problem by using a brutal grid search on different values of θ^*. At each value, the supply of skilled and unskilled workers are determined as in Appendix B.1. We then let the demand side of the markets run until they all clear. As explained in the main paper, the key difference between the planner's problem and the market equilibrium is that, wage premium in the planner's problem does not necessarily yield the θ^* picked by the planner.

As for ν, we follow the literature and try with values in a range of $[2, 5]$. The following figure shows that social welfare is well hump-shaped w.r.t. θ^*, at all values of ν. Moreover, the smaller ν is, the earlier welfare reaches its peak.

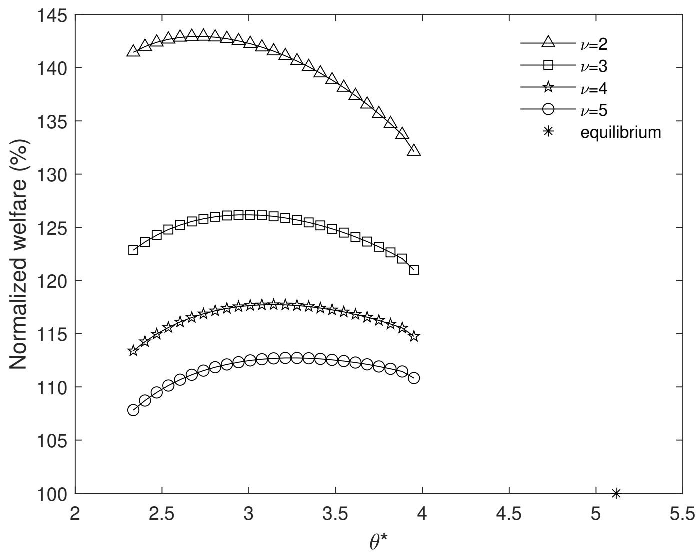


Figure D.1: Social Welfare as a Function of θ^*

Note: As in corresponding tables, we normalize the welfare of the competitive equilibrium to 100%, and use the $*$ dot near the right-end to represent it.

We also provide a table replicating results in Table 4.8. One can see that, the smaller ν is, or equivalently, the higher the I.E.S. is, the larger the welfare gap will be (last row). Which is intuitive, since a higher I.E.S. yields a larger welfare weight on the growth component g.

Table D.3: Comparison between Market Equilibrium and Social Planner's Allocation

Variable	**Benchmark**	**Planner**			
		$\nu = 2$	$\nu = 3$	$\nu = 4$	$\nu = 5$
θ^*	5.12	2.74	3.01	3.14	3.28
h^{supply}	3.71%	12.96%	10.74%	9.84%	9.05%
w^h/w^ℓ	2.43	1.28	1.44	1.52	1.59
g	1.97%	4.58%	4.34%	4.23%	4.13%
C_0	100%	94.57%	96.29%	96.91%	97.41%
U	100%	142.94%	126.17%	117.73%	112.72%

Note: The benchmark market equilibrium level of C_0 and U are normalized to 100%, respectively.

D.4 Effects of Generic R&D Tax Credit

The effects of the generic R&D tax credit is quite similar to what's with the quantity-based subsidy. They both result in more R&D, higher innovation intensity, but deteriorating innovation quality. The only difference is on the extensive margin: now entrants are investing less to become high-type, resulting a drop in p^*. In some sense, this policy is also "quantity-based" because it doesn't distinguish between R&D spending on innovations of different quality.

Table D.4: Strengthening the R&D Tax Credit

Variable	Meaning	Benchmark	$\mathbf{b_r}$+5%
$R(x)/Vadd$	average R&D intensity	15.75%	16.15%
$x_{\mathrm{H}d}/x_{\mathrm{H}}$	share of radical innovations	25.21%	24.13%
w^h/w^ℓ	wage premium	2.4318	2.4505
p^*	fraction of H-type entrants	18.94%	18.88%
g	TFP growth rate	1.97%	1.94%
C_0	initial consumption level	100%	99.76%
U	social welfare	100%	99.17%

Note: As in the main paper, x_{H} is calculated by aggregating $x_{\mathrm{H}}^{\mathrm{HTE}}$ and $x_{\mathrm{H}}^{\text{non-HTE}}$, same for $x_{\mathrm{H}d}$ and x_{L}. Innovation intensity is defined as the per line Poisson arrival rate of new innovations.

D.5 Demand Elasticity for Skilled Labor

We separate 2004 ASIE firms into two groups: those with positive R&D expenditure (innovating group) and those without (non-innovating group). Define *skilled labor ratio* of a firm as the fraction of total workers with college and above degrees. Wage premium is for each province and obtained from a Mincer type regression.[53]

For both groups, we run a regression of log skilled labor ratio against log provincial wage premium, controlling for size, SOE status, and industry dummies. The coefficient in front of log wage premium for the innovating group is -0.448, while that for the non-innovating group is only -0.156.

[53] We also tried to simply calculate the ratio of average wage of skilled labor to that of unskilled labor, the results are qualitatively the same.

We also compare the average skilled labor ratio between innovating and non-innovating firms in Table D.5. One can see that innovating firms typically hire more skilled labor.

Table D.5: Average Skilled Labor Ratio

	College	University	Graduate	No. of Firms
Firms with R&D exp.	18.22%	6.43%	0.39%	69,103
Firms without R&D exp.	9.45%	2.61%	0.16%	209,477
Firms with patents	20.58%	7.09%	0.48%	21,334
Firms without patents	10.83%	3.28%	0.19%	257,235
Firms with US-cited patents	26.62%	16.48%	2.81%	122
Firms without US-cited patents	32.32%	11.86%	1.17%	1,527