NBER WORKING PAPER SERIES

MILD GOVERNMENT FAILURE

Shang-Jin Wei
Jianhuan Xu
Ge Yin
Xiaobo Zhang

Working Paper 31178
http://www.nber.org/papers/w31178

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2023

We thank Daniel Xu, Nancy Qian, Ken Rogoff, Joe Stiglitz, Carlos Serrano and seminar and conference participants at the NBER Chinese Economy meeting, Columbia University, Emory University, World Bank, Asian Bureau of Financial and Economic Research (ABFER), and Asian Innovation and Entrepreneurship Association (AIEA) for their helpful comments and suggestions. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Mild Government Failure
Shang-Jin Wei, Jianhuan Xu, Ge Yin, and Xiaobo Zhang
NBER Working Paper No. 31178
April 2023
JEL No. H21,O14,O25

ABSTRACT

A relatively mild form of government failure - for example, bureaucrats can count but do not differentiate quality - can significantly affect the efficacy of industrial policy. We investigate this idea in the context of China's largest pro-innovation industrial policy using a structural model. We find that the return to the subsidy program is -19.7\% (but would be 7.8\% if the mild government failure can be removed). Furthermore, the welfare loss is exacerbated by patent trade.

Shang-Jin Wei
Columbia Business School
Columbia University
Kravis Building Room 552
665 West 130th Street
New York, New 10027
and NBER
shangjin.wei@columbia.edu

Jianhuan Xu
School of Economics
Singapore Management University
90 Stamford Road, Singapore
Singapore
xujh03@gmail.com

Ge Yin
Institute of International Economy
University of International Business and Economics
10 Huixin East Street
Beijing
China
yinge@uibe.edu.cn

Xiaobo Zhang
Guanghua School of Management
Peking University
and
International Food Policy Research Institute
2033 K Street, NW
Washington, DC 20006
x.zhang@gsm.pku.edu.cn

1 Introduction

When a market failure is identified, such as the existence of a positive spillover from innovation, a government intervention such as a subsidy to the innovating firms might raise the social welfare. Whether the intervention is warranted in the end depends on the relative importance of any government failure versus the market failure. While this principle is well understood in theory, the sign and the size of the net effect of the two forces in practice are subject to intense debates (see an extensive survey in Harrison and Rodríguez-Clare (2010)).

In this paper, we study the consequence of a relatively mild form of government failure - bureaucrats simply being average and not omniscient - on the success or failure of an industrial policy. For example, when a firm applying for a subsidy presents a set of recent patents as proof of its innovation ability to a government committee that reviews the application, the bureaucrats in the committee can count the patents but may not be able to differentiate their quality. The bureaucrats may not be less competent or more corrupt than their private-sector counterparts in this case as it is genuinely hard to differentiate the quality of a new innovation.[1] We show how this seemingly mild "deficiency" has an important effect on the efficiency consequence of a pro-innovation industrial policy. Precisely because this "deficiency" is mild and does not require the bureaucrats to be either more corrupt or less competent than other people, it is potentially more pervasive.

We study these questions in the context of China's largest pro-innovation industrial policy. The program is known as InnoCom and offers a large subsidy - a 10 percentage points reduction in the corporate income tax rate to successful applicant firms. A major policy change in 2008 expanded the scale of the program greatly, raising the fraction of eligible firms that can receive a subsidy from a negligible level to over 60%. Furthermore, it introduced a strong linkage between the success of a subsidy application and the number

[1] Drawing on a parallel from the empirical finance field which distinguishes among strong form, semi-strong form, and weak form of market efficiency, we distinguish among strong, semi-strong, and mild forms of government failure. A strong form of government failure is corruption by government officials or lobbying by powerful interest groups that cause government officials to deviate from maximizing social welfare (Rose-Ackerman (1975); Shleifer and Vishny (1993); Mauro (1995), Fisman and Wei (2004), and among others), which could compromise the efficiency of the industrial policy. A semi-strong form of government failure is the lower average competence of public sector employees than their private sector counterparts possibly due to lower relative public sector pay (Lazear (2000)). In comparison, a mild government failure refers to public sector decision-makers being just "average."

of recently granted patents that a firm owns.[2]. There are several reasons to study this program. First, given the international attention paid to China's innovation efforts, it is useful to understand the role of the country's industrial policy. Second, as many more countries are interested in promoting innovation using industrial policy, it is useful to have a framework to assess the relative importance of government failure versus market failure. The data from China's program provides a good angle to examine the issue.

The InnoCom program has some notable features (beyond its size). First, owning six new patents appears very important for applicant firms.[3] We will utilize this feature to identify some key model parameters. Second, externally purchased patents are counted in the same way as in-house innovations.[4] The existing literature on patent trade emphasizes its welfare-improving effects (see Serrano (2010) and Akcigit et al. (2016)). However, with distortions in the subsidy program, we show that patent trade can augment the welfare loss.

We develop four sets of salient empirical facts. First, we study how bureaucrats review subsidy applications. This uses a unique data set on the scores assigned by a government committee to the successful applicants for subsidy in a large Chinese city during 2008-2011 (the first four years of the new subsidy policy). We verify that the bureaucrats can count - they give a higher score to an applicant firm when its patent count goes up - but do not differentiate quality. In other words, we show that the mild form of government failure is present in our case rather than having to assume its existence. In addition, we detect that owning six new patents is especially important for applicant firms. Those firms with fewer than six new patents are less likely to receive a subsidy, but going beyond six does not materially raise the probability of success. We also verify that bureaucrats do not assign higher scores to firms with more in-house patents.

Second, we study the growth of patents following the 2008 policy shock. While the growth in the number of patents accelerated, the quality of patents declined significantly. We measure patent quality by patent renewal decisions by firms, forward citation count, and an estimated marginal contribution of a patent to firm productivity. All three proxies

[2]We use "patents" as a shorthand for "patents and other significant intellectual property rights (IPRs) such as sophisticated software that are registered with the National Software Bureau."

[3]"New patents" in the InnoCom program are defined as those granted within the previous three years.

[4]Allowing purchased patents to be used in subsidy applications is not unique to China. For example, a patent box policy common in Europe and Canada also has this feature.

reveal a notable decline in patent quality.

Third, we examine the behavior of the firms. In particular, the 2008 policy shock has induced the initially less innovative firms - those with fewer than six patents - in the targeted industries to rush to achieve the desired level of patents for subsidy applications. In addition, a rising share of the new patents owned by them appears to be of low quality. In comparison, other subsidy-eligible firms that start with six or more patents do not exhibit the same rate of patent growth.

Fourth, we study how the patent trade has changed following the 2008 policy shock. In particular, the share of patents sold to initially less innovative firms in the targeted industries exhibits the fastest growth after 2008. This is especially true for patents sold by either the firms outside the targeted industries, which are not eligible for a subsidy anyway, or by the firms in the targeted industries that already had more than six patents before the policy shock and hence do not need more to compete for a subsidy.

Inspired by these facts, we build a structural model to quantify the welfare impact of the subsidy program. The program affects welfare in several ways. On the plus side, an increase in high-quality patents can raise the productivity of all firms in the targeted industry through a positive spillover effect. However, it could also reduce the welfare through several channels. The first is direct resource waste from *subsidy-competing enterprises* (the SCEs) producing low-quality patents that may improve their chance of receiving a subsidy but otherwise do not raise productivity. The second is indirect resource waste from those firms not eligible for a subsidy but still engage in producing low-quality patents and selling them to the SCEs. The third is a new form of misallocation as high-value users of a patent that are not eligible for a subsidy may sell the patent to low-value users that are eligible for a subsidy. Last but not the least, because public funding is financed through distortionary taxation, it costs the society more than 1 RMB to fund 1 RMB worth of subsidy.

After calibrating the model to the data, we find that although the subsidy leads to an increase in the patent count by 33%, 98% of the increase is of low quality. This implies a notable decline in the average quality of the new patents. Since the program fails to induce a surge in high-quality patents, the positive gain from the additional productivity spillover is quite small.

On the other hand, as a low-quality patent is not too expensive to produce, the aggregate expenditure on low-quality patents is limited. In comparison, the social cost of public funding plays a large role in the ultimate welfare effect of the subsidy program. By comparing the welfare levels in the model with and without the subsidy program, we estimate the net social return to the subsidy to be -19.7%. That is, the society would be better off without this subsidy program.

In understanding the explosion of low-quality patents, we find that patent trade plays an important role. Without patent trade, the subsidy program would induce subsidy-eligible firms to produce low-quality patents. With patent trade, the program also inspires many firms not otherwise eligible for a subsidy, including those outside the targeted industry, to engage in the production of low-quality patents with the hope of selling them to subsidy-eligible firms. We show that a reduction in the frictions in patent trade by half would generate a further decline in the return to the subsidy program from -19.7% to -24%. This confirms a new "dark side" of patent trade.

We perform a number of counterfactual thought experiments. In particular, we can remove the mild government failure by allowing the bureaucrats to tell the quality of the patents and apply differential subsidy rules to high- and low-quality patents. In this case, the optimal policy would not subsidize low-quality patents. The return to the subsidy program would be 7.8%. Of course, removing the mild government failure is not feasible in practice. So the thought experiment serves to confirm that the presence of even a mild government failure could convert an otherwise well-justified industrial policy from success to failure. As another experiment, if the bureaucrats disallow the externally purchased patents in the subsidy applications, and adjust the subsidy on in-house patents optimally, the return to the subsidy program would have been 0.2%. This counterfactual may not be realistic either. If purchased patents were disallowed in subsidy applications, a market for "pre-patents" may emerge in parallel to a market for patents. That is, low-quality but patentable blueprints may be developed and sold by those firms not eligible for a subsidy to the subsidy-applying firms.

While we focus on mild government failure here, we certainly do not rule out strong or semi-strong forms of government failure in practice. If corruption, lobbying, or incompetence is incorporated into our model, the return to the subsidy program would have been even lower (i.e., more negative).

Our paper makes three contributions. First, we contribute to the literature on industrial policy (see the survey by Harrison and Rodríguez-Clare (2010)) by proposing the concept of a mild government failure. A key insight is that even a mild form of government failure could turn a theoretically sound industrial policy (i.e., one that corrects a well-identified market failure) into a government misadventure with a negative social return. This seems to be the first paper that demonstrates the existence of a mild government failure and documents its impact on the success or failure of a pro-innovation industrial policy.

Second, we contribute to the literature on patent trade by highlighting a new "dark side" of the trade. Most existing literature emphasizes that patent trade improves welfare (Shapiro (2010) and Lemley and Shapiro (2005)).[5] In comparison, we show that patent trade can also augment distortions through its interaction with mild government failure. In our case, lower frictions in patent trade reduce welfare. Our story is consistent with the "theory of the second best" (Lipsey and Lancaster (1956)). Our contribution is to demonstrate that this theoretical possibility has a practical bite in a large pro-innovation subsidy program.

Third, we contribute to the literature on China's innovation policy by highlighting important new channels of welfare losses from subsidy programs. Hu and Jefferson (2009) document a rapid rise of patents (and other innovations) in China. Wei et al. (2017) emphasize potential misallocation of innovation resources across firm ownership types. Chen et al. (2021) document the resource waste by firms re-labeling non-innovation expenditures as R&D expenditure. König et al. (2020) study input market frictions and how subsidy could induce some firms that should focus on imitation to sub-optimally switch to pursue innovation. Cao et al. (2022) document that China's subsidy programs seem to have induced a quantity-quality trade-off in the R&D investment (and they cited an earlier version of our paper on the evidence of a decline in patent quality). We point out several new channels that can reduce the efficiency of the subsidy program. They include the central role of a mild government failure. In addition, patent trade can amplify the loss. We also show in an extension that combining the distortion we study and the distortion associated with relabeling R&D expenditures would reduce further the return to the program.

The paper is organized as follows: In Section 2, we introduce the institutional back-

[5] An exception is research about "patent trolls" (Abrams et al. (2019)).

ground of the subsidy program and the data sets used in the paper. In Section 3, we document four sets of prominent data features. These data patterns motivate the setup of our model that is laid out in Section 4. Section 5 performs a number of counterfactual analyses. Finally, Section 6 concludes.

2 Background and Data

2.1 The Pro-Innovation Industrial Policy in China

China has several subsidy programs aiming at promoting innovation, and the largest such program is called InnoCom. According to the China Science and Technology Yearbook 2015, there are 16 pro-innovation programs administered by the Chinese central government, with a combined budgetary outlay of 154 billion RMBs. InnoCom is the largest program on the list. With an annual budget of 100 billion RMBs, it is much bigger than the sum of the other 15 programs.[6]

The InnoCom program aims to encourage innovation in what the Chinese government considers the "industries of the future". While the program started in the 1990s, a substantial change in the design in 2008 represents a major policy shock. First, the subsidy budget has grown substantially, raising the number of subsidized firms from a negligible share of the total number of firms that satisfy some basic conditions to over 60%. Second, the new approval process for the subsidy introduces an explicit linkage between the number of new innovations already owned by an applicant firm and the chance of obtaining a subsidy.

The program targets eight industries in advanced manufacturing and modern services.[7] Firms in the targeted industries have to first meet some threshold on R&D intensity (above 3%) and then have to show that they own a certain number of patents. Importantly, these patents can be acquired through patent trade and do not have to be developed in-house. This may be motivated by a desire by the architect of the program to broaden the set of firms that may be inspired to innovate. In other words, it may be

[6]For comparison, the US CHIPS and Science Act of 2022 allocates $52.7 billion of subsidy to US firms over 5 years, equivalent to about 72 billion RMBs a year using the exchange rate of 7 RMBs per dollar. This means that the annual expenditure of InnoCom in 2015 is bigger than the US program in 2022.

[7]They are pharmaceuticals (CSIC 27), special equipment manufacturing (CSIC 36), transportation equipment (CSIC 37), communication equipment and computers (CSIC 40), precision instruments (CSIC 41), computer service (CSIC 61), software service (CSIC 62), and environmental protection (CSIC 80).

considered desirable to be able to promote innovation by firms not applying for a subsidy, as long as these innovations are relevant to the targeted industries.

This feature is not unique to the Chinese pro-innovation industrial policy. The patent box policy - a pro-innovation subsidy program in the European Union, Australia, Britain, Canada, and other countries - also permits patents acquired by subsidy-eligible firms through patent trade to be used in qualifying for a subsidy. Gaessler et al. (2021) report that in two-thirds of their 15-country sample, patents acquired through patent trade are equally eligible for a patent box subsidy. Bösenberg and Egger (2017). Ciaramella (2017) find that a 1% increase in the tax subsidy in the patent box induces a 10% increase in patent trade. We have not found any study in the literature that suggests a distortionary consequence of this feature.

Unlike a patent box, InnoCom does not peg the subsidy to a portion of the profit that is self-reported by the firm to be linked to some patents. This may be motivated by a concern for potential arbitrariness in how a firm's profit is partitioned. [8] The InnoCom's requirement that only patents granted within the previous three years will be counted in the subsidy application review may be designed to encourage new innovation and avoid potential zombie patents that could be present in a patent box program.

In determining which firms will receive a subsidy, a government committee consisting of civil servants from the local tax bureau and the bureau of science and technology evaluates applicant firms by assigning numerical scores based on a count of patents, an ability to manage R&D, an ability to commercialize science and technology innovations, and growth potential. While the last three categories are subjectively assessed, the patent count appears objective. For the scoring purpose, one invention patent is considered equivalent to 6 utility or design patents. When the patents count increases, the applicant's score will increase as well until the patent count reaches six. Any higher count would not materially improve its chance of receiving the subsidy. A patent used in the application needs to be relatively new; the committee only counts those granted within the previous three years. On average, the program each year provides a subsidy to more than 4,600 firms in our sample city. At the end of 2010, 11,568 firms receive the subsidy, which is about 67% of all SCEs, or 20% of all firms in the eight targeted industries.

[8] As pointed out by (Griffith et al. (2014), partitioning a firm profit into a patent-related and a non-patent-related portions is subject to manipulation. Bloom et al. (2019) does not regard the patent box as a socially efficient tool to incentivize innovations.

We conjecture that the bureaucrats reviewing the applications for a subsidy can count the number of patents but are not able to tell their quality. Using our data on bureaucrats' scores on applicant firms in a city, we will confirm a lack of a positive association between the patent quality and the scores assigned by the bureaucrats. We will also confirm that bureaucrats do not value in-house patents more than externally purchased ones.

2.2 Data

While InnoCom is a national program, its implementation is carried out by local governments. For one large city, we have gained access to the scores assigned by the InnoCom committee to the successful applicant firms in the first 4 years of the program (2008-2011). This allows us to empirically check the roles of the patent count, patent quality, and patent origin (in-house development versus purchases from the patent market) in a firm's success in obtaining a subsidy.

To document the salient features of the data and calibrate the model, We utilize two additional data sets: administrative data about the firms from their tax records, and patent (and software) assignment data. The first data records firm-year level financial information (such as sales, employees, assets and etc.) and the industrial classification of the firms. The second data record the ownership of the innovations and any transfer of the ownership through trade. Using this information, we construct the patent portfolio for each firm each year.

For each patent, we know when it is first granted, whether it is renewed or not in each subsequent year when its ownership is changed, and who the buyer and the seller are. We compute its forward citation count. Unfortunately, we are not able to find information on ownership changes for software. But we will show that the broad picture of the count and quality of the innovations is driven by patents.

We use firm names to link up firm-level records in different data sets. We condense the firm names via the following steps. First, we keep only Chinese characters, letters, and numbers, and discard special symbols and punctuation marks. Second, we remove designations for corporate forms such as "limited corporation" or "subsidiary." Third, we convert all lowercase letters to upper cases. Two firm names in two data sets are considered the same if their names match after the above filters.

By this procedure, 94% of the firms that received an InnoCom subsidy in 2008 can be found in the patent assignment data too. This ratio seems satisfactory and is comparable to Hu and Jefferson (2009). In our sample, 90,539 patent (and software copyright) holders are found from 2005 to 2012. Financial information on firms is available from 2007 to 2011. We then separate the patents holders into two broad groups: *subsidy-competing-enterprises* (SCEs), which satisfy the basic eligibility of the InnoCom program, and *non-competing-entities* (NCEs), which are firms, individual innovators, or other institutions that are not eligible for the program. Each broad group can be further divided into two sub-groups as illustrated in Figure 1.

3 Salient Empirical Patterns

We document four salient data patterns related to the subsidy program, which will guide our subsequent structural model construction. Since InnoCom treats an invention patent as equivalent to six other patents, we convert every invention patent to six other patents in our analysis.

3.1 Bureaucrats can count but do not differentiate quality

Taking advantage of unique proprietary data on the evaluation scores made by the committee on all successful applicant firms that have received a subsidy in a large city in 2008, we study a number of questions.[9] First, can the bureaucrats tell the quality of patents? Second, do they care whether the patents are developed in-house by the applicant firms or purchased from other firms? Third, how much do they care about the number of patents?

We use two different proxies for patent quality: the renewal rate three years after the patent approval - which reflects the firms' own assessment of the usefulness of the patents, and the forward citation count three years after patent approval - which reflects the assessment by other inventors. As neither information is available to the bureaucrats at the time of a firm's subsidy application, we are examining if the bureaucrats have the ability to look for and analyze any soft information in the application process that helps them to forecast the quality of the patents. If they are able to do that, we would expect their scores on the applicants to be positively correlated with either the subsequent

[9]Unluckily, we only have this information for firms which were subsidized in 2008.

citation count or the renewable rate.

We regress the committee's points assigned to an applicant firm on a proxy for the quality of the patents portfolio, controlling for the firm's observed characteristics (sales, labor productivity, ownership, and industry) and its patent count.[10] In the first three regressions reported in Table 1, the dependent variable is the points that an applicant firm receives from the subsidy review committee. The first six regressors are dummies for patent count equaling one or two, three, four, five, six, and seven and above, respectively. The next regressor is a proxy for average quality of the patents owned by the applicant firm. The last regressor is the share of the patents that are developed in-house (as opposed to purchased through patent trade).

In column 1, we see that the firm's score tends to rise with more patents. The increases are big going from 4 or fewer to 5 patents, and from 5 to 6 patents. However, once the patent count reaches six, additional patents do not significantly raise the score. After controlling for the patent count, we see that the effect of higher patent quality in terms of the average subsequent citation count is statistically indifferent from zero. In column 2, we measure the patent quality by the renewal rates three years after the patent approval and find similar results. That is, bureaucrats do not systematically assign a higher score to those patents with a higher subsequent renewal rate. Since both forward citation count and patent renewal rate use future information, we are not saying that the bureaucrats should know them. Instead, the bureaucrats may have used all the soft and hard information available at the time of their decisions to try to gauge the innovation ability of the applicant firms. Our results confirm that they are not able to successfully differentiate the quality of the patents owned by the applicant firms. Again, since it is genuinely hard for anyone to tell the quality of patents ex ante, the bureaucrats are simply being average, and not omniscient.

In both columns 1 and 2, we also see that the share of in-house patents in total patents is not statistically significant. In other words, the bureaucrats do not assign more points to in-house patents than to those purchased from other firms. In column 3, we restrict the sample to those firms that do not use software in their applications. While this reduces the sample size to only 681 firms, all the results from the previous regressions stay the

[10]Collected by the Municipal Science & Technology Commission of a large city, the data includes the grading information of the four categories, as well as other firm characteristics.

same. In particular, once reaching six patents, additional patents do not significantly help with the scores. Holding the patent count constant, improvement in patent quality does not make a difference. Finally, there is no difference between purchased and internally developed patents.

In columns 4 to 6, we extend our sample to all SCEs and run a Probit equation

$$\begin{aligned} Pr(\text{Subsidy}_{i,t} = 1) &= \sum_n \beta_n \mathbb{1}(\text{patent count}_{i,t} = n) + \gamma \text{avg patent quality}_{i,t} \\ &\quad + X_{i,t} + \text{Year FE} + \epsilon_{i,t} \end{aligned} \tag{1}$$

where $Subsidy_{i,t}$ is a dummy variable that takes the value of one if applicant firm i in year t receives a subsidy, $X_{i,t}$ are observed characteristics of firm i in year t, including its industry classification, log sale and log TFP.[11] β_n and γ are the parameters of interested.

Again, we see no evidence of the bureaucrats' ability to tell patent quality as an applicant with higher quality patents receives no additional points. In turn, it will be reasonable for applicant firms to infer that the patent quality would not affect their chance of obtaining a subsidy.

Another interesting regressor is in-house patents as a share of an applicant firm's total patent counts. As the share of purchased patents is never significant, the bureaucrats do not seem to distinguish between in-house and externally purchased patents. This validates our interpretation that the subsidy program counts the number of patents that a firm owns, regardless of where they come from. When we restrict the sample to the firms that use only patents in their applications, we obtain the same results.

3.2 The patent quantity grows but the average quality declines

The growth of patents in China can be seen in the left graph of Figure 2, which plots the number of new patents including sophisticated software (blue-circle) and patents excluding software (red-square) per SCE by year. Compared to a linear trend using the data from 2004-2007 (depicted by the dashed lines), the growth accelerates after 2008. The acceleration is driven mostly by the patent growth since there is no divergence between the two lines.

[11] They are inferred from the firm's administrative tax data. Unfortunately, we do not observe the share of college workers and R&D workers for firms that do not receive a subsidy.

We gauge the patent quality in several ways. Our first measure is the fraction of the patents granted in a given year that are renewed three years later. Since renewing a patent requires an annual fee, the decision to renew reflects the owner's judgment on whether the patent is sufficiently valuable or not.[12]. The blue-circle line in the right graph of figure 2 plots the 3-years-out renewal rates of patents granted between 2004 to 2011. Before 2008, 87% of the patents are renewed three years after the initial patent approval. However, the renewal rate declines precipitously by 10 percentage points in 2008 and stays below a pre-2008 trend afterward.

Our second measure of patent quality is the average citation count three years after patent approval. The red-square line in the right graph of Figure 2 plots the average 3-years-out citation count as a function of the year in which the patents are granted. Similar to the renewal rate, we find that the forward citation count exhibits a sharp decline after 2008.

As a third measure of patent quality, we estimate the marginal contribution of an additional patent to a firm's productivity by regressing firm-level labor productivity (sales per employee) in a year on the number of newly obtained patents in the previous year, controlling for both industry and year fixed effects as well as firm location and ownership. To distinguish between the elasticity before and after 2008, we use a post-2008 dummy (inclusive) and interact it with the number of newly obtained patents. The result is reported in column 1 of table 2. We see a statistically significant decline in the marginal contribution of the new patents to firm productivity after 2008. Based on the point estimates in column 1, the labor productivity increases by 4.1% following an additional patent before 2007 but becomes negative (-6.0% + 4.1% =-1.9%) after 2008.

In column 2, we control for both firm fixed effects (which subsume the industry, location, and ownership dummies) and year fixed effects, and still find a similar result. In this case, while the marginal contribution of the patent ownership to firm productivity is no longer negative after 2008, it is still significantly smaller than before 2008. To summarize, while the patent quantity has grown tremendously after 2008, the quality has declined.

One can distinguish between inventions and non-invention patents. In column 3, we

[12]Using patent renewal as a proxy for patent quality is explored by Pakes (1986). A large literature follows this idea (such as Cornelli and Schankerman (1999), Lanjouw (1998), and Bessen (2008))

consider whether the drop in quality differs in these two groups of patents. The coefficient before the new invention is greater than the coefficient for non-invention patents. This suggests that the quality decline takes place for both invention and non-invention patents but the effect is more pronounced for non-inventions.

3.3 Less innovative SCEs simultaneously show a faster growth in the patent count and a larger decline in patent quality

As shown earlier, a prominent feature of the InnoCom program after 2008 is that the likelihood for a firm with six patents to receive a subsidy is much higher than those with fewer patents, but the average quality of the patents does not seem to matter much. Hence the policy shock may encourage less innovative firms - those initially with fewer than six patents - to dash for six patents. Figure 3 confirms this conjecture by presenting the density graphs of the patents owned by the SCEs in 2007 on the left and in 2008 on the right, respectively. In 2007, the year before the policy change, the density function is relatively smooth. The number of firms that own various number of patents decline monotonically with the number of patents. Importantly, there is nothing special about owning 6 patents. In contrast, as the InnoCom program was implemented at the beginning of 2008, the density of function suddenly exhibits a spike at 6 patents. This means that the firms clearly understand the importance of owning 6 patents. There is an unusually large number of firms that own exactly 6 patents, relative to those that own either 5 or 7 patents.

Inspired by the above figure, we examine the differences in behavior between the two groups of the SCEs with initially 6 or more patents versus with initially fewer than 6. We will label them as "initially more (or less) innovative SCEs", respectively. To compare the quantity of patents owned by them, we estimate the following equation

$$Y_{it} = \alpha_t \times D(patent_{it-1} < 6) + X_{it-1} + \mu_i + \mu_t + \epsilon_{it} \tag{2}$$

where Y_{it} is the patent count of SCE i in year t. $X_{it-1} = 1$ if firm i has obtained a subsidy within the previous 3 years, and 0 otherwise; μ_i and μ_t are firm and year fixed effects, which capture time-invariant firm heterogeneity and the aggregate trend, respectively. ϵ_{it} is an independent and identically distributed random error. α_t is the key parameter of interest, which measures the "excess" innovation count by the SCEs with initially fewer than six patents relative to those with initially six or more patents. For comparison, we do the same regression for the NCEs. Since these firms are not eligible to apply for an InnoCom subsidy, this can be regarded as a placebo.

The line of the blue circles in the left graph of Figure 4 traces out α_t from the SCE regression, representing the excess number of new patents by year by the initially less innovative SCEs over the more innovative SCEs. The capped spikes represent 90% confidence intervals.[13] The "excess" patent count is basically zero before 2008, implying no difference between them. However, α_t becomes significantly positive after 2008, suggesting that the initially less innovative SCEs dashed for more patents after the policy shock.[14]

As a placebo, the red squares in Figure 4 trace out the excess new patents by the less innovative NCEs over their more innovative counterparts. The difference between them is essentially zero for all years both before and after 2008. This suggests that the relative dash for more patents is a phenomenon especially prominent in the InnoCom-targeted industries.

We now check the evolution of relative patent quality. Let $V_{ikt} = 1$ be an indicator variable which equals to one if and only if SCE i chooses to renew patent k in year t. Let s_{ik} denote the year in which firm i obtains patent k. We consider the following regression:

$$V_{ikt} = \sum_{s \leq t} \beta_s \times D(patent_{is-1} < 6) + X_{it-1} + X_{kt} + \mu_i + \mu_t + \epsilon_{ikt} \tag{3}$$

where $D(patent_{is-1} < 6)$ is a dummy for firm i whose year s-1 patent count is strictly fewer than six. The summation means that for every year t, we consider all patents obtained before t. β_s is the key parameter of interest, representing the difference in the renewal probability by the initially less innovative SCEs relative to the more innovative ones. If the additional patents are acquired for the purpose of competing for a subsidy, rather than for their intrinsic productivity-enhancing value, then the new patents are likely to be of low quality and would not be worth the renewal cost once the InnoCom

[13]As we need the previous year's patent count to define the dummy D($patent_{it-1} < 6$), α_t starts from 2005.

[14]The difference between the two firm types becomes even stronger in 2011 possibly because it takes some time for all such firms to fully react to the policy shock.

review process is over. This would imply $\beta_s < 0$. On the other hand, if the firms acquire the patents for their true scientific value, then should not be a relative decline in the quality of the patents. In that case, $\beta_s = 0$.

Other controls are defined as follows. $X_{it-1} = 1$ if firm i has obtained a subsidy within the previous 3 years, and 0 otherwise. X_{kt} is a vector of observed characteristics of patent k, including the patient's age, industry classification, and type (invention, utility or design). μ_i and μ_t are firm and year fixed effects, respectively.

In the right graph of Figure 4, we plot β_s by patent vintage year owned by the SCEs (blue circles). We can see a clear relative decline in the patent quality after 2008 for the initially less innovative SCEs. This relative decline is not driven by any pre-trend. If anything, the average patent quality for the initially less innovative SCEs was higher in 2007 (the year before the policy shock). In the placebo test (red circles), the initially less innovative NCEs do not exhibit a relative decline in their patent quality relative to their initially more innovative counterparts.

3.4 Patent purchase by initially less innovative SCEs is the fastest growing type of patent trade

Since the InnoCom program accepts externally purchased patents when an SCE applies for a subsidy, it may be interesting to check if the initially less innovative SCEs show a disproportionately strong interest in purchasing patents than other firms after 2008. In Figure 5, we plot the shares of external patents in percent of all patents by the initially less innovative SCEs (with the solid blue circles), the initially more innovative SCEs (the red squares), and the NCEs (the light blue diamonds), respectively. For the initially less innovative SCEs, 2008 represents a structural break. Their share of purchased patents is doubled after 2008. In comparison, the initially more innovative SCEs only raise their share of purchased patents slightly after 2008, whereas the NCEs show no change around 2008.

There are interesting differences in the relative growth of patent trade for different seller-buyer combinations. In the left graph in Figure 6, we report the shares of traded patents sold to the SCEs with fewer than 6 initial patents from three other types of firms. The red-squre (the top line), blue-circle (the middle line), and green-diamond (the bottom line) represent the NCEs, the initially more innovative SCEs, and other less innovative

SCEs as sellers, respectively. The trade shares are stable before 2008. After the policy shock in 2008, the patent sales from both the NCEs and the more innovative SCEs to the less innovative SCEs jump up in percent of the total patent trade. In comparison, the patent sales from less innovative SCEs to other firms in the same category decline in relative importance. These patterns are entirely consistent with the incentive created by the 2008 policy shock.

The right graph in Figure 6 plots the shares of traded patents sold to either the initially more innovative SCEs or the NCEs from the initially less innovative SCEs (blue-circle line) and from the initially more innovative SCEs and the NCEs (red-square line). We see that both the trade among the initially more innovative SCEs and that within the NCEs decline sharply. In sum, the initially less innovative SCEs become major patent buyers following the 2008 policy shock, whereas other types of firms have raised their relative sales to the less innovative SCEs.

4 Model

We model the market equilibria with and without a subsidy program by incorporating the salient data patterns documented in the previous section. The subsidy program - in the style of the 2008 policy shock - counts the number of patents by applicant firms but does not adjust for patent quality, and does not distinguish between in-house versus purchased patents. We study how the program affects aggregate welfare by altering the firms' choices on patent production and trade. We consider a static model with three stages, denoted by $t \in \{1, 2, 3\}$, respectively.

4.1 Environment

There are two industries (the targeted industry and the non-targeted industry). A subset of the firms in the targeted industry satisfy the basic program requirement on the minimum share of R&D expenditure and are eligible to compete for a subsidy. They are referred to as Subsidy-Competing Enterprises (SCEs for short). There are two types of Non-subsidy Competing Entities (NCEs): those firms that are in the targeted industry but do not satisfy the minimum R&D requirement, plus those in the non-targeted industry. We use $i \in \{S, N_1, N_2\}$ to denote the SCEs and the two types of NCEs, respectively.

Firms differ in their initial patent count.

Patents are heterogeneous in their quality, denoted by x. For simplicity, we assume x is binomial with $x_H > x_L = 0$. Low-quality patents ($x = x_L$) have no intrinsic value and do not enhance the value of the firm (in the absence of a subsidy). As will be clear later, without a government subsidy, all patents in equilibrium are of high quality (x_H).

A firm can hold multiple patents. Let $n_{H,t}$ and $n_{L,t}$ denote the numbers of high- and low-quality patents that a firm holds in stage t, respectively. Since the bureaucrats do not distinguish patent quality, the subsidy decision depends only on the total number of patents in Stage 3, $n_3 = n_{H,3} + n_{L,3}$ but not on their composition. Denote the probability of receiving a subsidy for a firm of type i with n_3 patents by $\rho(i, n_3)$. While it is a policy choice variable, it is taken as given by the firms. As in the actual InnoCom program, that probability for the SECs, $\rho(S, n_3)$, is increasing in n_3 until it reaches $n_3 = 6$. On the other hand, as the NCEs are ineligible for a subsidy, $\rho(N1, n_3) = \rho(N2, n_3) = 0$.

After the subsidy decisions have been made and the production has taken place, each firm has to decide whether to renew the patents it owns at the cost of c per renewed patent. If a patent is renewed, it will bring a continuation value of $\pi x \varepsilon$, where π is a component of the profit that is determined by firm-specific productivity z. ε represents an obsolescence shock, i.e., the possibility that the patent may become useless in the future due to some other technological advances. We assume that $\varepsilon = 0$ with probability Ω_ε, and $\varepsilon = 1$ otherwise. Hence the additional value of holding a patent of quality x before the production is $\pi x + E_\varepsilon[\pi x \varepsilon - c]^+$, where $[.]^+$ equals 0 if the value inside the bracket is negative. The last term captures the firm's renewal decision. Since a given subsidy is good for one model period (or three years in the data), a patent's continuation value (and a firm's decision on patent renewal) does not depend on the subsidy.

We assume no complementarity among the patents. The expected value to the firm for holding $n_{H,3}$ patents is $\pi + n_{H,3}^{\alpha}(\pi x_H + E_\varepsilon[\pi x_H \varepsilon - c]^+)$. The marginal benefit of holding a high-quality patent is assumed to be decreasing in the number of such patents, i.e., $\alpha < 1$. As a result, the firm will only hold a finite number of patents. As low-quality patents generate no intrinsic benefit, they will be discarded once the subsidy decisions are made. Note that if $n_{H,3} = 0$, the firm's profit is just π. Hence π is the "autonomous profit" in the absence of a high-quality patent.

The autonomous profit of a firm with productivity z in the targeted industry (S

or N_1) is $\pi = Az$, where A is the industry-wide aggregate productivity. We assume $A = A_0 + \omega A_0 K^\eta$, where A_0 is the initial aggregate productivity, and K is the total number of high-quality patents in the targeted industry which will be determined endogenously in equilibrium. $\eta \geq 0$ and ω are parameters governing the knowledge spillover. Because A is taken as given by individual firms, the embedded spillover represents a well-defined market failure.

The autonomous profit of a firm in the non-targeted industry (N_2) is assumed to follow $\pi = z$, where the industry-wide productivity is normalized to 1. Those NCEs outside the targeted industry do not benefit from any knowledge spillover from the targeted industry. (This assumption is supported by the empirical findings in Table B1.)[15]

Firms also differ in their ability to produce high-quality patents, which is captured by a firm-type specific component of the innovation cost v. A firm's initial status can be summarized by its type $i \in \{S, N_1, N_2\}$, productivity z, innovation cost v, and the initial patent portfolio $(n_{H,1}, n_{L,1})$. We use $g(i, z, v, n_{H,1}, n_{L,1})$ to denote the measure of the firms with initial status $(i, z, v, n_{H,1}, n_{L,1})$.

In the first stage, a firm attempts to develop θ_H number of high-quality innovations and θ_L number of low-quality ones.[16] As it will be cleared later, in the absence of a subsidy, a low-quality idea does not generate any positive benefit and will not be patented by the firm. The innovation cost is denoted by $C(\theta_H, \theta_L; v)$ which is increasing and convex in both θ_H and θ_L.[17] In the second stage, a market in patent trade opens. A firm can choose to buy, sell or keep its patents. The price for a patent of quality x, $p(x)$, is determined competitively. In the last stage, an SCE with n_3 patents will receive a subsidy in the amount of $T\pi$ with probability $\rho(S, n_3)$. The subsidy is modeled as proportional to a firm's gross profit because the subsidy in the InnoCom program takes such a form. All firms then decide whether to renew each of their patents at a cost of c.

[15] For an individual inventor or a research institute, π can be interpreted as the utility of holding a patent.

[16] Note that n_H and n_L are a stock concept, referring to the numbers of high and low-quality patents, respectively, that a firm owns. In comparison, θ_H and θ_L are a flow concept, referring to the numbers of high and low-quality patents that a firm develops, respectively, in a given period.

[17] We allow the aggregate knowledge capital K to affect aggregate productivity but not the unit innovation cost. This is consistent with the usual assumption in the growth literature that the ratio of the R&D expenditures to GDP is not related to aggregate knowledge capital (Klette and Kortum (2004)) or at least does not decline in K (Akcigit et al. (2016)).

4.1.1 Production

We perform backward induction starting from the last stage. The value of a firm holding $n_3 = n_{H,3} + n_{L,3}$ number patents at the beginning of this stage is

$$V_3(i, z, n_{H,3}, n_{L,3}) = \pi + n_{H,3}^{\alpha} r(z) + \rho(i, n_3) T\pi$$

where $r(z) = \pi x_H + E_\varepsilon [\pi x_H \varepsilon - c]^+$ is the additional value from holding a patent with quality x. This implies a cutoff point for the patent renewal decision: if the continuation value is bad enough, the firm will not renew the patent and the additional benefit from the patent is 0.

4.1.2 Patent Trade

A firm with $n_{H,2}$ and $n_{L,2}$ number of high- and low-quality patents, respectively, before the trade can purchase or sell patents to maximize the gain from trade. Let m_H and m_L be the net purchases of high- and low-quality patents traded, respectively.

$$\begin{aligned} V_2(i, z, n_{H,2}, n_{L,2}) &= \max_{m_H, m_L} V_3(i, z, n_{H,2} + m_H, n_{L,2} + m_L) \\ &\quad + \rho(i, n_{H,2} + m_H + n_{L,2} + m_L) T\pi - p_H m_H - p_L m_L \end{aligned}$$

$$s.t. \ \ m_H \geq -n_{H,2} \text{ and } m_L \geq -n_{L,2}$$

Note that a negative value of either m_H or m_L indicates selling the patent. The two restrictions mean that the firm cannot sell more patents than what it owns.

A low-quality patent will have no buyers without the subsidy program. That is, if $\rho = 0$, $p(x_L) = 0$. Once the subsidy program is in place, then those SCEs with a low initial patent count may wish to purchase some. (Other SCEs with a high initial patent count would find no value in buying low-quality patents.) We can also see that the marginal befit to purchase patents is higher for large firms. Hence they will buy more patents and are more likely to be subsidized.

To allow for friction in market participation, we assume that a firm can participate in the patent trade with an exogenous probability σ. Hence the firm's value at the beginning of the second stage is $(1 - \sigma) V_3(i, z, n_{H,2}, n_{L,2}) + \sigma V_2(i, z, n_{H,2}, n_{L,2})$, where the first term uses the fact that $n_{H,2} = n_{H,3}$ and $n_{L,2} = n_{L,3}$ if the firm does not trade any patents.

4.1.3 Innovation stage

A firm with initial status $(i, z, v, n_{H,1}, n_{L,1})$ chooses the numbers of high- and low-quality innovations, θ_H and θ_L, respectively, to maximize its expected value $V_1(i, z, v, n_{H,1}, n_{L,1})$[18]:

$$\begin{aligned} V_1(i, z, v, n_{H,1}, n_{L,1}) &= \max_{\theta_H \geq 0,\ \theta_L \geq 0} (1-\sigma) V_3(i, z, n_{H,1} + \theta_H, n_{L,1} + \theta_L) + \\ &\quad \sigma V_2(i, z, n_{H,1} + \theta_H, n_{L,1} + \theta_L) - C(\theta_H, \theta_L; v) \end{aligned} \tag{4}$$

We impose a functional form on the innovation cost as

$$C(\theta_H, \theta_L; v) = \frac{v}{1+\zeta}\theta_H^{1+\zeta} + \frac{\bar{v}}{1+\zeta}\theta_L^{1+\zeta} + \frac{\hat{v}}{1+\zeta}\theta_L^{1+\zeta}\theta_H^{1+\zeta} \tag{5}$$

where the first and second terms are the R&D expenditures associated with the high and low-quality innovations, respectively. If $\hat{v}$ is positive, then the marginal cost of producing a high-quality innovation rises when more low-quality innovations are produced. (When we later calibrate the model parameters to match the moments in the data, $\hat{v}$ turns out to be a small positive number.)

We allow the cost parameter for producing high-quality innovation, v, to depend on both the firm type (whether it is an SCE, an NCE in the targeted industry, or an NCE outside the targeted industry) and whether the firm initially has any patent. These imply six separate parameters for the cost of producing high-quality patents, denoted by $v(SCE, n > 0)$, $v(SCE, n = 0)$, $v(NCE1, n > 0)$, $v(NCE1, n = 0)$, $v(NCE2, n > 0)$, and $v(NCE1, n = 0)$, respectively. In addition, two other cost parameters $\bar{v}$ and $\hat{v}$, are related to low-quality innovations, which are the same across all firms. In total, there are eight separate parameters describing the costs of innovations.

Without the subsidy program, since low-quality patents have no value, all firms would choose $\theta_L = 0$. This implies that before the subsidy program is in place, $n_{L,1} = 0$ as well. Since no firm pursues low-quality patents, we cannot identify $\bar{v}$ and $\hat{v}$ using only pre-InnoCom information. With the subsidy program in place, low-quality patents become

[18] In appendix A, we reformulate the model so that the representative firm chooses the number of projects x and the probability to convert a project to a high-quality patent q, instead of choosing the numbers of high- and low-quality patents directly. This means that the numbers of high-quality and low-quality patents are random variables. We show that our benchmark model is a first-order approximation of this model.

useful. Since the incentive to alter its innovation effort differs by firm type, the differential changes in patent renewal rates across different firm types will help identity $\bar{v}$ and $\hat{v}$.

Following the literature, we define aggregate knowledge capital as the sum of all R&D expenditure associated with high-quality innovations.[19]

$$K = \sum_{i \in \{S, N_1, N_2\}} \int \frac{v}{1+\zeta} \theta_H^{1+\zeta} g\left(i, z, v, n_{H,1}, n_{L,1}\right) dz dv dn_{H,1} dn_{L,1} \tag{6}$$

4.1.4 Welfare

We define the welfare level as the sum of all firm values net of the social cost of the subsidy, which can be written as follows:

$$Welfare = \sum_{i \in \{S, N_1, N_2\}} \int V_1\left(i, z, v, n_{H,1}, n_{L,1}\right) g\left(i, z, v, n_{H,1}, n_{L,1}\right) dz dv dn_{H,1} dn_{L,1} - (1+\tau) TS \tag{7}$$

where TS is the total subsidy. $\tau \geq 0$ denotes the marginal cost of collecting a dollar of tax. τ would be zero if there are neither distortions in the tax collection nor resource costs in administrating the subsidy program, In general, however, it costs society more than one dollar for every dollar of subsidy reaching a recipient firm.

The welfare effect of the subsidy program is the increase in the total firm profits minus the social cost of the subsidy. The social return to the subsidy is the change in the welfare due to the subsidy program in the percentage of the subsidy amount. The subsidy program may alter aggregate welfare in multiple ways. On the positive side, raising the number of high-quality innovations might generate a positive spillover to the productivity of all firms in the targeted industry. On the other hand, it could also reduce economic efficiency. First, those SCEs with a low initial patent count are incentivized to spend resources on low-quality innovations. Second, even firms not eligible for a subsidy may also spend resources on low-quality innovations with the hope to sell them to a subsidy-eligible SCE. Third, a high-quality patent may be sold by a higher-value user that is not eligible for a subsidy to a lower-value user that is eligible for a subsidy.

[19] The empirical work on the topic including Bloom et al. (2002) and Bloom et al. (2019) typically uses the total R&D expenditure to estimate the spillover from innovation. The endogenous growth literature (Akcigit et al. (2016)) also assumes that the spillover depends on the total R&D expenditure.

4.1.5 Characterizing the equilibrium

To characterize the equilibrium, we first ignore the integer constraint of the patent count. If the firm can trade, from the optimality conditions, we have

$$\alpha r(z)\, n_{H,3}^{\alpha-1} + \rho'(n_{H,3}+n_{L,3})\, T\pi = p_H \text{ and } n_{L,3} = (\rho' T\pi - p_L)^{+} \tag{8}$$

where ρ' is the marginal change in the probability of receiving a subsidy in response to a change in $n_{H,3}+n_{L,3}$. Hence for the NCEs, $n_{L,3}=0$. The SCEs, on the other hand, have the incentive to hold $n_{L,3}>0$ only if $n_{H,3}<6$. Meanwhile, in equilibrium, $\rho' T\pi \le p_L$; otherwise, the firm will increase $n_{L,3}$ until it reaches 6 and $\rho'=0$ in this case.

In the first stage, the optimality condition for θ_H yields

$$(1-\sigma)\left(\alpha r(z)(n_{H,1}+\theta_H)^{\alpha-1} + \rho'(n_{H,1}+\theta_H+\theta_L)\, T\pi\right) + \sigma p_H - \left(v + \hat{v}\theta_L^{1+\zeta}\right)\theta_H^{\zeta} = 0 \tag{9}$$

where the first two terms on the left-hand side mean that when a given patent is not sold, the marginal increase in the firm value is $\alpha r(z)(n_{H,1}+\theta_H)^{\alpha-1} + \rho' T\pi$. The third term is the marginal increase in the firm value when the patent is sold. The last term is the marginal cost of producing a high-quality patent.

Similarly, the optimality condition for θ_L yields

$$(1-\sigma)\,\rho'(n_{H,1}+\theta_H+\theta_L)\, T\pi + \sigma p_L - \left(\bar{v} + \hat{v}\theta_H^{1+\zeta}\right)\theta_L^{\zeta} = 0 \tag{10}$$

where we use $\frac{\partial V_2}{\partial \theta_L} = p_L$. The SCEs have a stronger incentive than NCEs to produce θ_L since SCEs' $\rho' \ge 0$. For any firm type, the incentive to produce low-quality patents also depends on $\hat{v}$.

In the case of no subsidy (i.e., $T=0$), we have $\theta_L = n_{L,3} = 0$ and

$$(1-\sigma)\,\alpha r(z)(n_{H,1}+\theta_H)^{\alpha-1} + \sigma p_H = v\theta_H^{\zeta} \tag{11}$$

$$n_{H,3} = \left(\frac{\alpha r(z)}{p_H}\right)^{\frac{1}{1-\alpha}} \tag{12}$$

The two sides of equation 11 are the marginal benefit and marginal cost of generating a new patent, respectively. By Equation 12, when the firm has a chance to trade, it always adjusts the number of patents it owns so that the marginal benefit of a patent is p_H.

We summarize the welfare effect of the subsidy as follows. Once the subsidy program is in place, low-quality patents become valuable in the equilibrium, $p_L > 0$. From equation (9), firms in the innovation stage waste some resources to produce low-quality patents that do not generate any positive spillover. This would push up the marginal cost of producing high-quality patents and hence potentially reduce the amount of positive spillover. On the other hand, the subsidy program also increases the benefit of high-quality patents. These two opposing forces jointly determine the net change in the knowledge spillover.

With the possibility of engaging in patent trade, a firm would always choose the number of patents in such a way that the marginal value of a patent equals its price in the patent market (as in equation (8)). As long as $\rho' > 0$, the subsidy creates a wedge between the marginal production value of a patent and its market price, which can induce misallocation (i.e., a patent is bought by a firm eligible for a subsidy from another firm that would have been its highest-value user. In the production stage, the subsidy will involve an opportunity cost τ.

4.2 Calibration

We calibrate the model to match some key moments in the data. Since each subsidy approval is good for three years, we set one period in the model to be equal to three years in the data. We treat the period of 2005-2007 as a period with no subsidy ($\rho = T = 0$) [20] Once the subsidy program is introduced in the model economy, we match $\rho(i, n)$, the probability to receive a subsidy conditional on both the firm type and the patent count, to be exactly those estimated in Column 5 of Table 1. We set $T = 0.1$ to capture the form of the InnoCom subsidy (a reduction in the corporate income tax by 10 percentage points for three years). We normalize the targeted industry's pre-shock aggregate productivity, A_0, to be 1. Its value after the policy shock will be determined endogenously. We use a firm's actual patent count at the end of 2004 to represent a firm's initial patent count in the model in the absence of the subsidy program.

[20] Very few firms received a subsidy during 2005-2007. In any case, the limited amount of subsidy was not tied to an applicant firm's patent count or other innovation outcomes.

4.2.1 Parameters set with prior information

A firm's productivity conditional on its type i is assumed to follow a log-normal distribution, with mean μ_i and standard deviation Ω_i, respectively, where subscript i denotes whether a firm is an SCE, an NCE in the targeted industry, or an NCE outside the targeted industry. We estimate these six parameters from the actual distribution of the firms' profits for each firm type. Interestingly, we find that, on average, the SCEs are more productive than the NCEs in the targeted industry ($\mu(N1) = 0.83\mu(S)$), which in turn are more productive than the NCEs outside the targeted industry ($\mu(N2) = 0.62\mu(S)$). We find that the Ω's are approximately the same across the firm types.

To pin down the value of a high-quality patent, x_H, we recognize that, without the subsidy program, no low-quality patents are produced. Since the increase in a firm's profit (before 2008) due to a patent is πx_H, we use the average profit increase due to a new patent to infer x_H, which we estimate to be 3.4%.[21]

The knowledge spillover parameter, η, is important for determining the gains from the subsidy program. It is estimated by Bloom et al. (2002)) to range from 0 to 0.2. To err on the side of giving a more favorable interpretation of the subsidy program, we set $\eta = 0.2$ in the baseline case. We will later perform robustness checks for different values of the parameter. In Appendix B, we also report an estimate based on our data (which turns out to be smaller than 0.2).

The shadow value of the public funds τ is estimated by Chen et al. (2021) and Ming (2009) to be between 0.2 and 0.4. We set $\tau = 0.2$, the lower bound of their estimate in the baseline case. This would also give the most favorable interpretation of the subsidy program.

For the scaling parameter in the innovation cost function, ζ, we follow Peters (2020) and set it to be one. For parameter α that governs the diminishing return to firms holding more patents, we follow the "span of control" literature (Lucas 1978) and set $\alpha = 0.7$.

[21] In the data, denote $\tilde{\pi}_t$ as the profit of a firm in year t before 2008, and n_t is the number of new patents obtained in t (either in-house or external patents). Since a model period is 3 years, we have that $\frac{3\Delta \ln \tilde{\pi}_t}{n_t} = x_H$.

4.2.2 Parameters jointly calibrated to data moments

There are 12 additional parameters in the model. First, there are 8 parameters describing the costs of innovation, including 6 with regard to producing high-quality innovations ($v\,(SCE, n = 0)$, $v\,(SCE, n > 0)$, $v\,(NCE1, n = 0)$, $v\,(NCE1, n > 0)$, $v\,(NCE2, n = 0)$, and $v\,(NCE2, n > 0)$), one with regard to producing low-quality patents, $\hat{v}$, and a remaining one with regard to the interaction between producing high- and low-quality patents, $\bar{v}$. Second, 2 parameters affect patent renewal decision: the probability of an obsolescence shock, Ω_ε, and per-period cost of renewing a patent, c. Third, σ the probability that a firm can participate in patent trade reflects the friction in the patent market. Fourth, ω is a scaling parameter converting high-quality knowledge to industry-level productivity.

These 12 parameters are jointly calibrated by minimizing the sum of the distances between a set of model moments and the corresponding data moments. In particular, We target the following 12 moments in the data: the average new patent count per each firm type during 2005-2007 conditional on whether there is any patents in 2004, the share of externally purchased patents traded before 2008, the targeted industry profit growth rate before the subsidy, the 3-years-out renewal rate of the patents granted before 2008 by firms with labor productivity below and above the median, respectively, and the relative differences in the patent count and renewal rate, respectively, between the SCEs with and without 6 initial patents before and after 2008. These targeted data moments and their values are listed in Table 4.

Although all parameters are calibrated together, it may be useful to discuss intuitively which variations in the data may play an important role in identifying which parameters in the model. Note that before the policy shock, all patents can be regarded as of high quality. Consider first Ω_ε and c. Since a firm will renew a patent if $\pi x_H \varepsilon > c$, the observed renewal decision on an in-house patent by a firm conditional on its productivity helps to identify Ω_ε and c. The observed share of the external patent helps to identify σ. Finally, the six parameters associated with the cost of producing high-quality patents by firm type are identified based on the number of innovations by firm type before the subsidy program.

For the two parameters associated with producing low-quality patents, $\bar{v}$ and $\hat{v}$, we use a difference-in-differences regression. We allow for possible coincidental changes in factors other than the subsidy program and assume that such changes affect the intrinsically more and less innovative SCEs (those with six or more initial patents and those without) in the same way. We interpret the result from the double differences (between the more or less intrinsically innovative firms and before and after the policy shock) as reflecting the effect of the subsidy program. We then simulate the model one more time by imposing $\rho(i, n_3)$ and $T = 0.1$ after the policy shock, while keeping other parameters unchanged. Since those intrinsically less innovative SCEs (with fewer than 6 initial patents) have a stronger incentive to produce low-quality patents, the double differencing result can identify $\bar{v}$ and $\hat{v}$.

4.2.3 Parameter values and model fitness

We summarize all parameter values in Table 3 and report the model fitness in Table 4. It is unsurprising that, for the 12 targeted moments reported in the top panel, the model moments match the data well. Meanwhile, we check four important but untargeted "double relative" moments reported in the lower part of the table, including the patent count, external patent share, and renewal rate. All of them are the differences in the changes between SCEs and NCEs before and after the policy shock. The model fitness appears reasonably good as well.

Figure 8 plots the density of patents held by the SCEs after trade with and without the subsidy. Consistent with the data pattern in Figure 3, we can see that most SCEs in the model hold just 6 patents. As a unit of value in the model is calibrated to be 1 million RMBs, the renewal cost, $c = 0.01$, implies that the renewal cost per patent per year is about 3,000 RMB (0.01mil/3). This cost includes not only what is paid to the patent registration office (about 1,000 RMB), but also legal fees and administrative costs to cover activities needed to safeguard against infringement.

It is interesting to observe that the estimated innovation cost of high-quality patents, v, is smaller for those firms with initial patents than those firms without. This cost differential is consistent with the interpretation that those with initial patents have a stronger intrinsic ability to innovate. It is also interesting to note that the cost differential is smaller for the SCEs than for the NCEs. This is not surprising since only firms with a high enough R&D expenditure can become SCEs. With $\bar{v} = 0.003$, the cost to produce a low-quality patent is low (about 3,000 RMB). On the other hand, with $\hat{v} = 0.01\%$, more low-quality patents raise the unit production cost of a high-quality patent.

Since the model only targets the difference in the changes in patent counts between the SCEs with low and high patent counts, we use the difference in the changes in the shares of external patents between these two types of SCEs as a validation test of our model. In the data, the difference in the share of purchased patents between the two types of SCEs rises by 13%. In the model, this difference rises by 7%, which goes in the same direction. Meanwhile, we also use the difference between patent renewal rates, the count of newly invented patents, as well as external shares between NCEs with six and more initial patents and below six initial patents as another untargeted moment. In the data, all these differences between the two types of NCEs only change a little, which is similar to the model prediction. Overall, the model does a reasonably good job at matching these untargeted moments as well as the targeted ones.

5 Assessing the Subsidy Program

The Quantity and Quality of Patents

To understand the effect of the subsidy program, we compare the quantity and quality of patents by firm types with and without the subsidy and report the results in Table 5. We separate the firms into four types: the SCEs with initial patents below 6, the SCEs with initial patents above 6, the NCEs in targeted industries, and other NCEs. In the first column, we report the patent count (in thousands) held by each firm type before the subsidy. Because there are more SCEs with a low initial patent count than those with a high count, the former group collectively owns more patents.

In columns 2 to 4, we report the patent count by firm type in a laissez-faire economy (i.e., no subsidy or $T = 0$). Not surprisingly, no firm wants to produce or own low-quality patents in this economy. Column 2 reports the number of patents owned before the trade, which is equal to the initial patent count (column 1) plus the number of newly developed patents. We see that those SCEs with a high initial patent count invent more because they have a lower invention cost.

The net purchase of patents by firm type is reported in Column 4. A negative number means that a particular firm type is a net seller. Note that the sum of net purchases across all firm types is zero. The SCEs, especially those with a high initial patent count, are net sellers, whereas the NCEs are net buyers. This suggests that, in the absence of the subsidy program, the SCEs have a comparative advantage in producing high-quality patents.

What happens to patent production and trade in an economy with the subsidy program (i.e., $T = 0.1$) is described in the next three columns. Those SCEs with a low initial patent count now buy a lot of patents. Interestingly, as reported in Column 7, they buy substantially more low-quality patents (about 11,700) than they do high-quality ones (only 1,200). Because low-quality patents do not enhance productivity, the purchase of low-quality patents is motivated by a desire to exploit the mild government failure - that bureaucrats count patents but do not differential quality - for a chance to receive a subsidy.

All other types of firms are net sellers of low-quality patents. This is very telling: as none of them has any intrinsic use for low-quality patents, these firms produce them only with the hope to sell them to those SCEs who need them to apply for a subsidy. Given the friction in the patent trade market, many of the producers of low-quality patents do not succeed in selling all of them.

In the last column, we characterize the subsidy economy in terms of its differences from the laissez-faire economy. While high-quality patents go up slightly (by 0.9 thousand), low-quality patents increase by a substantially greater amount (by 48.2 thousand). In other words, 98.1% of the increased quantity of the patents due to the subsidy program is of the low-quality variety.

The Social Return to the Subsidy Program

We infer the social return to the subsidy program by computing the percentage change in the welfare level (the sum of all firms' profits net of the social cost of the subsidy) from the laissez-faire economy to the subsidy economy. The details are reported in Table 6. In particular, we calibrate the model twice, reporting the outcomes for each firm type in the laissez-faire and subsidy economies in Columns A and B, respectively. The monetary unit is a billion RMBs.

In Panel (1), we report R&D expenditures on high- and low-quality patents in the first stage. In terms of the expenditure on high-quality patents, we see that those SCEs with a low initial patent count are the only type of firms that have been induced by the subsidy program to spend more on high-quality patents. This is understandable since other types of firms do not need to do more than what they already find optimal under laissez-faire.

In comparison, in the subsidy regime, every type of firm finds it optimal to start to produce low-quality patents (also see column (5) in the previous table). While those SCEs with a low initial patent count hope to use the low-quality patents to raise their chance of receiving a subsidy, other types of firms produce low-quality patents with the hope of selling them to the former group. Note that the overall increase in the R&D expenditure stimulated by the subsidy program is not too high because the low-quality patents are not too expensive to produce. (In an appendix, we will show a bigger increase in the reported R&D expenditures due to mislabeling of some of the non-R&D expenditures.)

In Panel (2), we report the net purchase expenditures from patent trade for each firm type. Going from laissez-faire to the subsidy regime, we see that those SCEs with a low initial patent count have switched their stats from net sellers of patents to net buyers. In comparison, all other firm types have either increased their net sales (in the case of those SCEs with a high initial patent count) or reduced their net buys (in the case of the two types of NCEs). These patterns in terms of monetary values are of course consistent with net purchase quantities in the previous table. It is also interesting to note that the price of a low-quality patent has increased dramatically from zero under laissez-faire to 3000 RMBs under the subsidy regime. In comparison, the price of a high-quality patent has increased by only slightly (i.e., less than 1/3 of a 1% from 350 to 351 thousand RMBs).

In Panel (3), we report the revenues from production and the subsidy. Because all firms in the targeted industry benefit from a spillover due to a (modest) increase in the number of high-quality patents, the revenues of both the SCEs and those NCEs in the targeted industry also go up (by a modest amount). The revenue for the NCEs outside the targeted industry declines slightly because the high-quality patents they buy have become slightly more expensive.

Given the sums of the firm profits under laissez-faire and the subsidy regime, respectively (reported in Row 4), the increase in the economy-wide firm profits per period for our sample city before subtracting the social cost of the subsidy is 8.68 billion RMBs (Row 5). Recall that the subsidy probabilities differ by firm type and are calibrated based on the actual data from our sample city. The corresponding total social cost of the subsidy is 10.38 billion RMBs per period (Row 6) [22]. This means that the social return to the pro-innovation subsidy program is -19.7% (Row 7). In other words, this industrial policy is not a success in spite of a modest increase in the spillover from high-quality patents.

In summary, this is a setting with a well-identified market failure - the existence of a positive spillover to the industry-wide productivity from firm-level innovation. However, our structural model shows that the subsidy program has mostly inspired the production of more low-quality patents that do not have a positive productivity effect and very few additional high-quality patents. Given the social cost of the subsidy, this industrial policy produces a negative return.

Correspondence to the Main Data Patterns

It may be useful to comment on how the model setup and the predictions are related to the four salient data patterns documented in the previous section. Recall that the first empirical pattern in Section 3 is that the bureaucrats who review the firm applications count the number of patents but do not differentiate their quality. This is directly built into the subsidy rule in the model. In particular, for an SCE with a given number of patents in the model, the probability that it may receive a subsidy is directly matched to the probability estimated in column 5 in Table 1.

The second set of data patterns is that the subsidy program appears to have generated simultaneously an increase in the quantity of patents but a decline in their average quality. The model generates this pair of results as described in Table 5. In particular, going from laissez-faire to the subsidy regime, the new patent count has increased by 33% (from 150,700 to 199,800 reported in the second to the last row). However, out of the newly increased patents, 98.1% are of low-quality variety. This implies a large decline in the average quality of the new patents produced.

The third set of data patterns is that the less innovative SCEs simultaneously show faster growth in the patent count but a larger decline in patent quality than their more innovative counterparts. The model incorporates this pattern as a targeted moment as reported in Table 4. In particular, from the last two targeted moments, we see that after

[22]Recall this uses the lower bound estimate of the social cost of public funding in China - 1.20 RMB to the society for every 1 RMB of the subsidy received by the firms. If one were to use a median estimate, the social return to the subsidy would be lower.

the subsidy, those SCEs with fewer than six initial patterns increase their low-quality patents much more than the other SCEs with more initial patents. The same patterns can be discerned in Tables 5 and 6 as well.

A comparison between more or less innovative NCEs (which do not directly compete for a subsidy) is akin to a placebo test since their relative incentive to produce patents should not be altered as much by the subsidy program. In our model, from the first two un-targeted moments, the difference in the change in the patent count following the subsidy program between those NCEs with fewer than six initial patents and those with more than six patents is not significantly different from each other. This replicates the empirical pattern in Figure 4.

The fourth set of data patterns is that less innovative SCEs become the major patent buyers of patents after the subsidy program (Figures 5 and 6). Our model also produces these patterns, as can be seen from the last two un-targeted moments in Table 4. (The same conclusion can be drawn from Tables 5 and 6 as well.) In particular, the SCEs with fewer initial patents in the model purchase significantly more patents than other SCEs following the subsidy program. In contrast, the gap in the patent purchase behavior between the less and more innovative NCEs is not affected much by the subsidy program. In sum, our structural model appears to have captured the essence of all the data patterns that are documented in Section 3.

Sensitivity Checks

We perform two types of robustness checks. We start by examining the sensitivity of both the model parameters and the return to the subsidy to small perturbations in the data moments. In panel A of Table 7, the first column lists all the data moments used to generate the parameters in the model. In each row, we increase a given moment in the data by 5% and then report the percentage change in each internally calibrated parameter relative to the initial level in Table 6. The first six data moments are the counts of the new patents invented over 2005-2007 by each firm type with and without any initial patent in 2004. They help to identify the six corresponding cost parameters in the model for producing high-quality patents.

For example, the first data moment is the number of patents invented over 2005-2007 by those SCEs without any patent in 2004. It is the key data moment that helps to identify the cost of producing high-quality patents in the model for that type of firm. If this data moment is raised by 5%, our calibration would reduce the cost parameter, v(S, n=0), of producing high-quality patents by that type of firm by 9.9%. Since all model parameters are calibrated jointly, it is not surprising that some other parameters also change, albeit by a much smaller magnitude. Similarly, when each other data moment (new patent count during 2005-2007) for a particular firm type is raised by 5%, we can see that the cost parameter in the model for that corresponding firm type also declines by close to 10%. It is important to note that a small increase in any of these six data moments, in the end, leads to no significant change in the estimated return to the subsidy program (as reported in the last column). This suggests that our conclusion regarding the efficiency consequence of the subsidy program is not sensitive to possible measurement errors in these data moments.

The 7th and 8th data moments in the first column are the two average patent renewal rates by all firms and the firms with their intrinsic productivity component z above the median, respectively, during 2005-2007. They are the key data that help to identify the cost of renewing a patent, c, and the probability of an obsolescence shock, Ω_ϵ. Unsurprisingly, the renewal cost c in the model declines when the observed renewal rate in the data goes up by 5%. For the highly productive firms - those with z greater than the median value, their patent renewal decision depends only on Ω_ϵ. A higher observed renewal rate in the data implies a lower likelihood of an obsolescence shock. Most importantly, from the results reported in the last column, we see that the estimated return to the subsidy program is not sensitive to a small change in these two data moments either.

The 9th data moment - the average share of the patents held by all firms during 2005-2007 that were purchased from the patent trade - is the key data point that helps to quantify the friction in the patent trade market. An increase in this data moment by 5% implies a higher probability of patent trade in the model, σ, by 1.8% (which means less friction to patent trade). Such a change in the data would reduce the ultimate estimate of the return to the subsidy program by a tiny bit (from -19.7% to -19.8%).

Finally, the last two data moments are the changes in the patent renewal rate and new patent counts, respectively, by the less innovative SCEs (those with fewer than 6 initial patents) due to the subsidy program relative to those by the more innovative SCEs. As each is essentially a double-differenced outcome, we assign a "relative-relative" label to them. These data moments are used to identify the cost parameters associated with producing low-quality patents. When the relative-relative renewal rate in the data is increased by 5%, $\bar{v}$ in the model rises by 1.8%. The direction of the change is not surprising: A relative-relative renewable rate in the data means that even the less innovative SCEs must face a reduced incentive to produce low-quality patents, and a slightly higher cost of producing low-quality patents in the model can produce such a result. From the last column, we see that the estimated return to the subsidy program is barely changed.

An increase by 5% in the observed relative growth in patent count by the less innovative SCEs after the subsidy program translates into a reduction in $\bar{v}$ and $\hat{v}$ by 32.7% and 11.4%, respectively. The sign of these changes is also intuitive. If the less innovative SCEs could manage to increase their relative patent count by a greater amount, it must imply that the cost of producing low-quality patents is even lower than before. This leads to a very small deterioration in the estimated return to the subsidy program.

To take stock, we note that the estimated return to the subsidy after each perturbation to the data moments is always very close to the benchmark case, i.e., between -19.1% and 20.9%. These results indicate that the model parameters and, most importantly, our conclusion on the efficiency consequence of the subsidy program is not very sensitive to either small measurement errors in the data or small perturbations to the data moments for other reasons.

In Panel B of Table 6, we conduct a different type of robustness check. In particular, we increase the internally calibrated parameters by 5% one by one and re-compute the return to the subsidy program in each case. For example, if $v(S, n = 0)$ in column 1 is increased by 5% (from 0.02) while holding all other model parameters constant, the estimated return to the subsidy program would become -19.9%. This represents a very small change from the baseline estimate of -19.7%.

We do similar comparative statics for each of the other calibrated parameters in the other columns. From the results reported in the last row, we can see that the estimated return to the subsidy is within a relatively narrow range between -19.1% and -20.0%. In other words, our conclusion regarding the return to the subsidy program is not very sensitive to a small change in any of the model parameters.

The Central Role of the Mild Government Failure

Consider a thought experiment in which mild government failure does not exist. In particular, the bureaucrats can distinguish the patent quality and choose different probabilities of granting a subsidy depending on the patent quality. In this case, it is optimal to only subsidize the high-quality patents. Let T_H be the tax cut for an SEC with six high-quality patents. We search for the optimal values of T_H that would maximize the welfare while holding all other model parameters as given. We find that $T_H = 0.15$ would maximize the welfare. In other words, the optimal program design would not count low-quality patents but would subsidize firms with a high-quality patent with even a bigger tax cut (15 percentage points reduction in the tax rate) than in the current subsidy program (10 percentage points reduction).[23]

This result is reported in Column 1 of Table 8. As we see from Panel (1.2), in the absence of a subsidy for low-quality innovations, no firm would waste resources on them. On the other hand, comparing Panel (1.1) of this column with the corresponding numbers in Column A of Table 6, all firms now invest more in producing high-quality patents. In contrast to the baseline case (Column B of Table 6), we now see an increase in the number of patents with no deterioration in their average quality. Patent trade now plays a positive role in raising social welfare: those SCEs with a low initial patent count now buy more high-quality patents from other firms (Panel (2) in the first column). The positive spillover to productivity from the increased high-quality patents now raises the revenues of all three types of firms in the targeted industry (as seen in Panel (3) in the first column).

Given these changes induced by the subsidy, the social return to the subsidy program, in this case, is 7.8%. The positive social return is not surprising since there is a well-defined market failure in this economy - the existence of a positive spillover to industry-level productivity from firm-level high-quality innovations that have not been internalized by individual firms. In the absence of mild government failure, the subsidy to high-quality innovations mitigates market failure.

Note that we are not saying that real-world bureaucrats should be able to differentiate

[23]We have also examined a case in which potentially separate subsidies can be given to in-house versus externally purchased patents. We find that, in the absence of mild government failure, the optimal policy would not distinguish between in-house versus purchased high-quality patents, and would give a 15 percentage points tax cut in both cases (but no tax cut to low-quality patents).

the quality of patents as it is intrinsically hard to do so. Instead, this thought experiment makes the point that the existence or absence of a mild government failure can have a significant consequence on the efficacy of an industrial policy. While we intentionally assume away a strong form or a semi-strong form of government failure in the current paper, we are not saying that corruption, lobbying, or incompetence is irrelevant to the implementation of industrial policies. If we are to add these ingredients to the model, the social return to the subsidy program would be even lower (i.e., more negative).

The Role of Patent Trade

To examine the interactions between patent trade and industrial policy, we consider three thought experiments. The first is to reduce the friction in patent trade by doubling the probability that firms can participate in the patent market, σ, from 33 percent to 66 percent. The results are reported in columns 2 (a laissez-faire economy) and 3 (the subsidy regime) of Table 8.

In the subsidy regime (relative to the baseline case of more friction to patent trade), the total subsidy expenditure is increased to 8.98 billion RMBs (as more SCEs would now be able to buy and own a patent). While the number of patents increases, the quality of innovation declines further, as indicated by a decline in the renewal rate. This translates into an even lower return to the subsidy (-24%) than the benchmark case (-19.7%) in Table 6. Relative to the existing literature on patent trade (Akcigit et al. (2016)) that highlights the benefits of patent trade in improving resource allocation, our case illustrates the possibility that patent trade could worsen resource misallocation in the presence of a mild government failure.

As a second experiment, we consider a case in which only in-house patents can be used in the subsidy application (while maintaining the values of all other parameters and σ=0.33). That is, SCEs cannot rely on patent trade to boost their chance of receiving a subsidy. The results are reported in column 4 of Table 8. This change in the design of the subsidy program removes the incentive to produce low-quality patents by all firms other than those directly applying for a subsidy. This reduces the quantity of the low-quality patents produced as well as the expenditures on them.

Without the ability to use purchased patents to apply for a subsidy, fewer SCEs would apply for a subsidy. Compared to the baseline subsidy program described in column B of

Table 6), the total subsidy expenditure declines to 8.32 billion. These savings represent a substantial improvement over the baseline case. As a result, the social return to the subsidy is now -10% (rather than -19.7%). Note that the return to the subsidy is still negative because those SCEs with a low initial patent still have the incentive to waste resources on producing low-quality patents. In fact, they try harder to do so in order to compensate for not being able to use purchased patents in their subsidy applications.

In the third experiment reported in column 5, we fix the total subsidy budget at the baseline level (8.65 billion RMBs) while excluding purchased patents from the subsidy consideration. This implies a bigger tax cut per firm. Relative to the original program, the new program also delivers a better outcome (with the return to the subsidy at -15%). An important part of the improvement comes from avoiding resource waste in producing low-quality patents by firms not directly applying for a subsidy. However, relative to the previous case of not fixing the subsidy budget, the subsidy is too large for the size of the productivity spillover, and it exacerbates the incentive for the SCEs to produce low-quality in-house patents.

Optimal Policy Subject to Mild Government Failure

A possible improvement in the design of the program is to assign different levels of subsidies to in-house and externally purchased patents (while the bureaucrats still cannot differentiate the quality of patents). We solve for their optimal values numerically while taking as given all other model parameters. We find that the optimal tax cut is 12% for the in-house patents and 0% for the external patents. In other words, the optimal policy should not count the externally purchased patents in the subsidy decisions. Moreover, in the presence of mild government failure, the optimal policy also needs to substantially reduce the subsidy for in-house patents. The results from this experiment are reported in Column 6 of Table 8. The total subsidy expenditure declines to 0.76 billion RMBs. The social return to the optimal subsidy program is now 0.2%.

It is useful to compare this case to the "first-best" case in Column 1 of Table 8. Without mild government failure, the optimal policy does not need to distinguish between in-house versus purchased patents, but will only subsidize high-quality innovations. Indeed, we find that the optimal subsidy is 15 percentage points reduction in the tax rate, which is 50% larger than China's actual industrial policy. Since it is not realistic to remove the mild government failure, the "first best" policy described above is not a feasible policy.

It may be tempting to regard the policy described here as a constrained optimum. We note that distinguishing between in-house versus purchased innovations may not be straightforward either. If the subsidy rule only allows for in-house patents, the subsidy applicants may have an incentive to buy "pre-patents" from other firms - i.e., patentable innovations to be disguised as in-house innovations. For example, a black market for "pre-patents" may emerge that would match a potential buyer - an SCE in need of a patent - with a prospective seller - before an external innovation is patented. Then it is a question of how costly it is for the bureaucrats to tell whether a given innovation is truly developed in-house. The cost may be especially high in some developing countries where bureaucratic incompetence or corruption is severe. In such economies, our results suggest that the pro-innovation industrial policy is more likely to fail.

Size of the Spillover

A key parameter governing the importance of market failure is the size of productivity spillover from innovations. The estimates of the spillover parameter, η, by Bloom et al. (2002)), based mostly on data from advanced economies, range from 0 to 0.2. To err on the side of giving a more favorable interpretation of the subsidy program, we set $\eta = 0.2$ in the baseline case. It is possible that the actual size of spillover is smaller for a country like China since it is not on the technological frontier. To assess the sensitivity of the conclusion, we plot the returns to the subsidy as a function of η in Figure 9. The solid blue line indicates the return to the subsidy as a function of η when $\tau = 0.2$. As we can see, for all realistic values of η, the subsidy program always yields a negative return. In particular, if the spillover parameter for China is 0.10, i.e., which is in the middle of the empirically estimated range (between zero and 20%), the return to the subsidy program would be -90%. The dashed line represents the return to the subsidy when only in-house patents are counted in the subsidy applications. As we see, the return would be somewhat better than when externally purchased patents are included, though still negative.

For an easy comparison, we plot the case of no mild government failure in the dashed-dot orange line. The return to the subsidy program (where the subsidy to high-quality innovation is optimally computed) will always be positive for η between 0.1 and 0.2. Of course, the greater the spillover, the higher the return to the subsidy program. When

$\eta = 0.20$, the social return to the subsidy program is 7.8%, exactly as reported at the bottom of Column 1 in Table 8.

In Appendix B, we estimate the size of the spillover using Chinese data and find it to be around 0.01, which is much smaller than the one used in our baseline calibrations. This would imply a much lower return to the subsidy program.

Relabeling R&D expenditures

Some NCEs in the targeted industry may pretend to be SCEs by relabelling some non-R&D expenditures (such as management costs) as R&D expenditures. This is carefully studied by Chen et al. (2021), who estimates the average relabelling cost to be 24% of the true R&D cost. We now extend our model to include this distortion. Each N_1 firm (in the targeted industry) can now choose to pretend to be an SCE at the beginning of the innovation stage. The overall cost inclusive of manipulation and innovation is $C(\theta_H, \theta_L; v) + q$, where q is a fixed cost to relabel. We calibrate relabeling cost q to be 24% of the true R&D (following Chen et al. (2021)). It is easy to verify that if $\rho = 0$, no firm would want to incur the extra relabelling cost. We compare two policy designs: the baseline subsidy program and an alternative policy that only counts in-house patents. In both cases, we keep ρ unchanged.

Table 9 reports the result. As we can see, with possible relabeling, the subsidy in both cases increases substantially, as more firms become eligible for a subsidy. When purchased patents are allowed in the subsidy application, the true R&D cost (excluding the relabelling cost) is lower than the benchmark case. For the NCEs to pretend to be SCEs, the innovation cost rises. By diverting financial and labor resources to relabelling, they reduce the true R&D resources. Thus, the aggregate productivity declines, and the return on subsidy is -26%, 7 percentage points lower than the case of no R&D mislabeling.

We now exclude externally purchased patents in the subsidy application. In column 2, the total subsidy declines to 20.17 billion RMBs. The social return to the subsidy is -7%, better than allowing purchased patents, but still worse than without the R&D mislabeling. To summarize, even with the possibility of relabeling R&D expenditure (so that some of the subsidy-eligible firms are fake SCEs), the social returns to the subsidy program are better when externally purchased patents are excluded for consideration in the subsidy applications.

6 Conclusion

We study how even a mild form of government failure - the bureaucrats can count but do not differentiate quality - affects the efficacy of a large pro-innovation industrial policy in China. We show that the presence or absence of this feature could change the sign on the return to the subsidy program. In particular, without the mild government failure, the return to the program would be 7.8%. In contrast, with mild failure (which is a more realistic scenario), the return to the subsidy program is -19.7%.

We also pay special attention to the role of patent trade. By counting purchased patents by applicant firms in competing for a subsidy, the architect of the industrial policy may hope to use patent trade to inspire more innovations even by firms not directly competing for a subsidy. (This feature is not unique to Chinese industrial policy, and is shared by the patent box policy in Europe and elsewhere). Our estimation suggests that this feature lowers the return to the subsidy program. We also show a reduction in the friction to patent trade could exacerbate the welfare loss of the program in the presence of mild government failure.

References

Abrams, D. S., U. Akcigit, G. Oz, and J. G. Pearce (2019). The patent troll: Benign middleman or stick-up artist? Technical report, National Bureau of Economic Research.

Akcigit, U., M. A. Celik, and J. Greenwood (2016). Buy, keep, or sell: Economic growth and the market for ideas. *Econometrica 84*(3), 943–984.

Bessen, J. (2008). The value of us patents by owner and patent characteristics. *Research Policy 37*(5), 932–945.

Bloom, N., R. Griffith, and J. Van Reenen (2002). Do r&d tax credits work? evidence from a panel of countries 1979–1997. *Journal of Public Economics 85*(1), 1–31.

Bloom, N., J. Van Reenen, and H. Williams (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives 33*(3), 163–84.

Bösenberg, S. and P. H. Egger (2017). R&d tax incentives and the emergence and trade of ideas. *Economic policy 32*(89), 39–80.

Cao, L., H. Jiang, G. Li, and L. Zhu (2022). Haste makes waste? quantity-based subsidies under heterogeneous innovations.

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

Ciaramella, L. (2017). Patent boxes and the relocation of intellectual property. Technical report.

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

Fisman, R. and S.-J. Wei (2004). Tax rates and tax evasion: evidence from "missing imports" in china. *Journal of political Economy 112*(2), 471–496.

Gaessler, F., B. H. Hall, and D. Harhoff (2021). Should there be lower taxes on patent income? *Research Policy 50*(1), 104129.

Griffith, R., H. Miller, and M. O'Connell (2014). Ownership of intellectual property and corporate taxation. *Journal of Public Economics 112*, 12–23.

Harrison, A. and A. Rodríguez-Clare (2010). Trade, foreign investment, and industrial policy for developing countries. *Handbook of development economics 5*, 4039–4214.

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

Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Journal of political economy 112*(5), 986–1018.

König, M., Z. M. Song, K. Storesletten, and F. Zilibotti (2020). From imitation to innovation: Where is all that chinese r&d going? Technical report, National Bureau of Economic Research.

Lanjouw, J. O. (1998). Patent protection in the shadow of infringement: Simulation estimations of patent value. *The Review of Economic Studies 65*(4), 671–710.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review 90*(5), 1346–1361.

Lemley, M. A. and C. Shapiro (2005). Probabilistic patents. *Journal of Economic Perspectives 19*(2), 75–98.

Lipsey, R. G. and K. Lancaster (1956). The general theory of second best. *The review of economic studies 24*(1), 11–32.

Lucking, B., N. Bloom, and J. Van Reenen (2019). Have r&d spillovers declined in the 21st century? *Fiscal Studies 40*(4), 561–590.

Mauro, P. (1995). Corruption and growth. *The quarterly journal of economics 110*(3), 681–712.

Ming, L. (2009). Estimation and analysis of marginal cost of public funds in china [j]. *Collected Essays on Finance and Economics 6*.

Pakes, A. (1986). Patents as options: Some estimates of the value of holding european patent stocks. *Econometrica*, 755–784.

Peters, M. (2020). Heterogeneous markups, growth, and endogenous misallocation. *Econometrica 88*(5), 2037–2073.

Rose-Ackerman, S. (1975). The economics of corruption. *Journal of public economics 4*(2), 187–203.

Serrano, C. J. (2010). The dynamics of the transfer and renewal of patents. *The RAND Journal of Economics 41*(4), 686–708.

Shapiro, C. (2010). Injunctions, hold-up, and patent royalties. *American Law and Economics Review 12*(2), 280–318.

Shleifer, A. and R. W. Vishny (1993). Corruption. *The quarterly journal of economics 108*(3), 599–617.

Wei, S.-J., Z. Xie, and X. Zhang (2017). From“ made in china” to“ innovated in china”: Necessity, prospect, and challenges. *Journal of Economic Perspectives 31*(1), 49–70.

Tables and Figures

Table 1: How are applicants scored?

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Points (OLS)			Subsidy or Not (Probit)		
	Avg citation	Avg renewal rate		Avg citation	Avg renewal rate	
Patent count=1 or 2	0.000	0.000	0.000	0.062	0.063	0.043
	(0.000)	(0.000)	(0.000)	(0.075)	(0.076)	(0.118)
Patent count=3	0.945**	0.912**	-1.037	0.147	0.148	0.025
	(0.418)	(0.335)	(1.526)	(0.084)	(0.086)	(0.131)
Patent count=4	0.965**	0.914*	1.809	0.167***	0.169***	0.057
	(0.442)	(0.547)	(1.674)	(0.026)	(0.029)	(0.065)
Patent count=5	3.131***	3.133***	3.040***	0.238**	0.239**	0.128
	(0.419)	(0.856)	(1.964)	(0.085)	(0.087)	(0.121)
Patent count=6	6.643***	6.590***	6.624***	0.319***	0.311**	0.198*
	(0.417)	(0.551)	(1.871)	(0.069)	(0.071)	(0.107)
Patent count> 6	7.693***	7.614***	7.452	0.353***	0.354***	0.203*
	(0.499)	(0.473)	(1.697)	(0.046)	(0.048)	(0.101)
Quality proxy	0.007	-0.727	-0.459	0.007	0.011	0.007
	(0.007)	(0.473)	(1.697)	(0.016)	(0.009)	(0.103)
Share of In-house Patents	-0.007	-0.007	-0.006	0.001	0.001	-0.005
	(0.004)	(0.006)	(0.008)	(0.013)	(0.014)	(0.015)
Share of college workers	1.062	1.095	0.563			
	(0.769)	(0.761)	(1.613)			
Share of R&D workers	-0.205	-0.264	0.825			
	(0.628)	(0.752)	(0.926)			
ln(sale)	1.307***	1.302***	1.279***	0.003	0.003	0.002
	(0.145)	(0.131)	(0.190)	(0.002)	(0.002)	(0.002)
ln(TFP)	-0.712***	-0.704***	-0.596***	0.002	0.001	0.001
	(0.159)	(0.102)	(0.187)	(0.001)	(0.001)	(0.001)
Industry FE	Y	Y	Y	Y	Y	Y
Year FE	-	-	-	Y	Y	Y
Only firms without software			Y			Y
Obs.	2,470	2,470	791	7,166	7,166	5,289
Adj. R2	0.17	0.17	0.22	0.16	0.16	0.15

Notes: This table shows the correlation between 2008 subsidized firms' patents quality and their evaluation by the bureaucrat. The dependent variables are total points evaluated by the bureaucrat in columns 1 to 3, while in columns 4 to 6, the dependent variables are dummy variables which equal 1 if the firm first gets the subsidy and 0 otherwise. Quality proxy is the average citation (columns 1 and 4) or renewal rates (other columns) of patents owned by a firm. In column 3 and 6, the sample is restricted to firms without software. In-house patent share is the share of self-developed patents. All standard errors are clustered at the industry level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Labor Productivity and Patents

	ln(labor prod$_{it}$)	ln(labor prod$_{it}$)	ln(labor prod$_{it}$)
New patent$_{it-1}$	0.041***	0.043**	
	(0.013)	(0.025)	
(New patent$_{it-1}$) $\times$ ($t \geq 2008$)	-0.060***	-0.037*	
	(0.021)	(0.022)	
New invention$_{it-1}$			0.058**
			(0.030)
(New invention$_{it-1}$) $\times$ ($t \geq 2008$)			-0.025**
			(0.009)
New non-invention$_{it-1}$			0.011
			(0.010)
(New non-invention$_{it-1}$) $\times$ ($t \geq 2008$)			-0.056**
			(0.027)
Year FEs	Y	Y	Y
Firm FEs		Y	Y
Obs.	31,332	10,995	10,995
Adj. R2	0.42	0.77	0.77

Notes: This table shows the marginal change of firm labor productivity when the patent count increases. New invention and New non-invention are the counts of new obtained inventions and non-inventions. All standard errors are clustered at the firm-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Parameter Values Used in the Baseline Simulation

Parameter	Value	Description	Source
$v(S, n = 0)$	0.04	SCE ($n = 0$): high quality patent innov. cost	Calibration
$v(N1, n = 0)$	1.57	NCE1 ($n = 0$): high quality patent innov. cost	Calibration
$v(N2, n = 0)$	0.75	NCE2 ($n = 0$): high quality patent innov. cost	Calibration
$v(S, n > 0)$	0.02	SCE ($n > 0$): high quality patent innov. cost	Calibration
$v(N1, n > 0)$	0.04	NCE1 ($n > 0$): high quality patent innov. cost	Calibration
$v(N2, n > 0)$	0.01	NCE2 ($n > 0$): high quality patent innov. cost	Calibration
$\bar{v}$	0.003	Low quality patent innovation cost	Calibration
$\hat{v}$	0.0001	Low quality patent innovation cost	Calibration
σ	0.33	Probability of participating in patent trade	Calibration
c	0.01	Renewal cost	Calibration
Ω_ϵ	0.11	Prob of obsolescence shock	Calibration
ω	7.99	Level of knowledge spillover	Calibration
ζ	1.00	Curvature of innovation cost	Acemoglu et al. (2018)
α	0.70	Span of control	Lucas (1978)
x_H	0.03	Value of high quality patent	Estimated from data
η	0.20	Elasticity of knowledge spillover	Lucking et al. (2019)
τ	0.20	Marginal shadow cost of 1 RMB public funding	Chen et al. (2021)

Notes: The model value is 1 million RMB.

Table 4: Model Fitness

Targeted Moments	Data	Model
Pre-subsidy new patent count		
by firms with no initial patent:		
- SCEs	3.18	3.21
- NCEs in targeted industries	2.10	2.08
- NCEs outside targeted industries	0.36	0.36
Pre-subsidy new patent count		
by firms with some initial patents:		
- SCEs	3.60	3.52
- NCEs in targeted industries	2.16	2.13
- NCEs outside targeted industries	0.48	0.47
Pre-subsidy external patent share	0.03	0.03
Pre-subsidy patent renewal rate:		
- All firms	0.83	0.83
- Firms with above median productivity	0.89	0.89
Pre-subsidy profit growth rate	0.03	0.03
Relative difference in new patent count due to subsidy		
between intrinsically less and more innovative SCEs	1.80	1.80
Relative difference in patent renewal rate due to subsidy		
between intrinsically less and more innovative SCEs	-0.13	-0.15
Non-targeted Moments		
Relative difference in renewal rates due to subsidy		
between intrinsically less and more innovative NCEs	0.01	0.03
Relative difference in patent counts due to subsidy		
between intrinsically less and more innovative NCEs	0.03	0.01
Relative difference in shares of purchased patents due to subsidy		
between intrinsically less and more innovative SCEs	0.13	0.07
Relative difference in shares of purchased patents due to subsidy		
between intrinsically less and more innovative NCEs	0.00	0.00

Notes: This table reports the model's fit by comparing the moments in the model and the data. The phrase "less and more innovative" firms compares those firms with fewer than six initial patents relative to those with six or more initial patents.

Table 5: Patent Quality and Quantity Before and After the Subsidy Program (unit=1,000 patents)

		Laissez Faire			Under Subsidy			Δ patents from subsidy
	(1) Initial patents	(2) pre-trade	(3) post-trade	(4) net buy	(5) pre-trade	(6) post-trade	(7) net buy	(8)=(6)-(3)
Total high quality patents (a)	68.1	150.7	**150.7**		151.6	**151.6**		**0.9**
- SCEs with initial patents < 6	12.4	21.5	20.4	-1.1	21.8	23.0	1.2	2.6
- SCEs with initial patents ≥ 6	9.5	32.7	21.6	-11.1	33.3	20.9	-12.4	-0.6
- NCEs in targeted industry	20.4	53.5	54.2	0.6	53.5	54.0	0.5	-0.2
- NCEs outside targeted industry	25.8	42.9	54.6	11.6	42.9	53.7	10.7	-0.9
Total low quality patents (b)	0.0	0.0	**0.0**		48.2	**48.2**		**48.2**
- SCEs with initial patents < 6	0.0	0.0	0.0	0.0	13.6	25.3	11.7	25.3
- SCEs with initial patents ≥ 6	0.0	0.0	0.0	0.0	0.4	0.0	-0.4	0.0
- NCEs in targeted industry	0.0	0.0	0.0	0.0	8.0	5.4	-2.6	5.4
- NCEs outside targetted industry	0.0	0.0	0.0	0.0	26.2	17.5	-8.6	17.5
Total number of patents(c= a+b)	68.1	150.7	**150.7**		199.8	**199.8**		**49.1**
Share of low quality patents (d = b/c in%)	0.0	0.0	0.0	-	24.1	24.1	-	**98.1**

Notes: This table reports the patent quantity and quality for each firm type in two scenarios: laissez faire (no subsidy) and the economy with subsidy. The unit is 1,000 patents except for the last row.

Table 6: Expenditures and Profits (in billion RMBs) and the Return to Subsidy (%)

	(A) Laissez Faire	(B) With Subsidy
(1) Innovation Expenditure	**6.15**	**6.33**
(1.1) R&D cost for high quality patents	6.15	6.23
- SCEs with initial patents < 6	0.65	0.82
- SCEs with initial patents ≥ 6	1.68	1.59
- NCEs in targeted industry	2.44	2.44
- NCEs in non-targeted industry	1.38	1.38
(1.2) R&D cost for low-quality patents	0.00	0.10
- SCEs with initial patents < 6	0.00	0.08
- SCEs with initial patents ≥ 6	0.00	0.003
- NCEs in targeted industry	0.00	0.01
- NCEs in non-targeted industry	0.00	0.01
(2) Net Purchases from Patent Trade	**0.00**	**0.00**
- SCEs with initial patents < 6	-0.38	0.46
- SCEs with initial patents ≥ 6	-3.87	-4.66
- NCEs in targeted industry	0.22	0.21
- NCEs in non-targeted industry	4.04	3.99
Price of high-quality patent p_H (in 1,000 RMBs)	*350*	*351*
Price of low-quality patent p_L (in 1,000 RMBs)	*0*	*3*
(3) Revenue from Production and Subsidy	**271.89**	**280.75**
(3.1) Output Value Excluding Subsidy	271.89	272.10
- SCEs with initial patents < 6	14.45	14.49
- SCEs with initial patents ≥ 6	37.21	37.32
- NCEs in targeted industry	77.24	77.33
- NCEs in non-targeted industry	143.00	142.97
(3.2) Subsidy	0.00	8.65
- *Average ρ among SCEs*	*0%*	*67%*
(4) Total Firm Profit Inclusive of Subsidy		
= (3) -(1)-(2)	**265.74**	**274.42**
(5) Increase in Total Firm Value (=(4B)-(4A))		**8.68**
(6) Social Cost of the Subsidy (=(1+τ)*Subsidy)		**10.38**
(7) Return to the Subsidy (= ((5)-(6))/Subsidy)		**-19.7%**

Notes: This table reports the economy in two scenarios: without subsidy (column 1) and with subsidy (column 2). The unit is billion RMBs except for p_H and p_L which are in thousand RMBs.

Table 7: Sensitivity Checks

	Panel A: % Parameter changes to 5% increase in a data moment												Ret. to sub (%)
	$v(S,n=0)$	$v(N1,n=0)$	$v(N2,n=0)$	$v(S,n>0)$	$v(N1,n>0)$	$v(N2,n>0)$	$\bar{v}$	$\hat{v}$	σ	c	Ω_ϵ	ω	
Data moments \ Parameter values	0.04	1.57	0.75	0.02	0.04	0.01	0.003	0.0001	0.33	0.01	0.11	7.99	-19.7
Responding to 5% increase in each data moment													
Patent # by firms w/o initial patent													
- SCEs	-9.9	0.0	-0.1	-0.2	0.1	0.1	0.0	0.0	0.0	-	-	-0.6	-20.1
- NCEs in targeted industry	0.0	-9.8	0.0	-0.3	0.1	0.0	0.0	0.0	0.0	-	-	-1.0	-19.5
- NCEs in non-targeted industry	0.0	0.0	-9.8	-0.3	0.1	0.0	0.0	0.0	0.0	-	-	0.0	-19.8
Patent # by firms with initial patent													
- SCEs	0.0	0.0	0.0	-9.8	0.1	0.0	0.0	0.0	0.0	-	-	-1.5	-19.7
- NCEs in targeted industry	0.0	0.0	0.0	-0.3	-9.4	0.0	0.0	0.0	0.0	-	-	-1.3	-19.8
- NCEs in non-targeted industry	0.0	0.0	0.0	-0.3	0.1	-9.4	0.0	0.0	0.0	-	-	0.0	-19.8
Renewal rates w/o subsidy													
- All firms	0.3	0.5	0.8	0.2	0.6	0.5	0.0	0.0	0.0	-44.2	0.0	-0.4	-19.1
- Firms with $z >$ median	1.6	1.8	1.6	1.7	1.9	1.7	0.0	0.0	0.0	38.3	-40.5	-2.0	-19.7
External patent share w/o subsidy	0.4	0.7	0.6	0.5	0.8	0.7	0.0	0.0	1.3	-	-	-0.4	-19.8
Relative patent renewal rate	-0.2	0.0	0.0	-0.3	0.1	0.0	1.8	0.0	0.0	-	-	0.0	-19.6
Relative patent count	0.0	0.0	0.0	0.0	0.0	0.0	-32.7	-11.4	0.0	-	-	0.0	-20.9
	Panel B: Sensitivity of the return to subsidy to 5% increase in each parameter												
Return to subsidy (%)	-19.9	-19.6	-20.0	-19.5	-19.9	-19.9	-19.7	-19.1	-19.7	-19.7	-19.7	-19.2	

Notes: Panel A reports the change in the calibrated parameters and the return to subsidy when a targeted data moment increases by 5%. Panel B reports the return to subsidy when each parameter increases by 5%. The far left column reports a set of targeted data moments. "Relative patent renewal rate"

Table 8: Alternative Subsidy Policies

	No mild govmnt failure	$\sigma = 0.64$		Only count In-house patents		Optimal subsidy with mild govnt failure
		Laissez Faire	With subsidy	$T = 0.1$	with a fixed budget	
	(1)	(2)	(3)	(4)	(5)	(6)
(1) Innovation Expenditure	6.67	9.00	9.38	6.47	6.45	6.22
(1.1) R&D cost for x_H	6.67	9.00	9.02	6.20	6.20	6.18
- SCEs with initial patents < 6	0.94	0.97	1.19	0.67	0.67	0.67
- SCEs with initial patents ≥ 6	1.87	2.52	2.32	1.68	1.68	1.69
- NCE in targeted industry	2.46	3.62	3.62	2.46	2.46	2.44
- Other NCE	1.39	1.89	1.89	1.39	1.39	1.38
(1.2) R&D cost for x_L	0.00	0.00	0.36	0.25	0.27	0.04
- SCEs with initial patents < 6	0.00	0.00	0.21	0.25	0.27	0.04
- SCEs with initial patents ≥ 6	0.00	0.00	0.01	0.00	0.00	0.00
- NCE in targeted industry	0.00	0.00	0.04	0.00	0.00	0.00
- Other NCE	0.00	0.00	0.10	0.00	0.00	0.00
(2) Value from Trade	0.00	0.00	0.00	0.00	0.00	0.00
- SCEs with initial patents < 6	0.23	-2.27	0.75	0.14	0.14	0.02
- SCEs with initial patents ≥ 6	-4.23	-7.19	-10.80	-4.08	-4.07	-4.21
- NCE in targeted industry	0.19	1.20	1.21	0.14	0.14	0.19
- Other NCE	3.81	8.26	8.84	3.80	3.79	4.00
Price of high-quality patent p_H (in 1,000 RMBs)	*340*	*340*	*340*	*350*	*350*	*340*
Price of low-Quality patent p_H (in 1,000 RMBs)	*0*	*0*	*3*	*0*	*0*	*0*
(3) Revenue from Production and Subsidy	276.92	295.97	304.99	281.34	281.29	272.87
(3.1) Output value excluding subsidy	273.39	295.97	296.01	273.02	272.64	272.11
- SCEs with initial patents < 6	14.25	15.61	16.16	14.74	14.64	14.01
- SCEs with initial patents ≥ 6	38.52	38.21	37.70	37.89	37.63	37.89
- NCE in targeted industry	77.84	88.02	88.02	77.65	77.65	77.26
- Other NCE	142.78	154.13	154.13	142.74	142.72	142.95
(3.2) Subsidy	3.53	0.00	8.98	8.32	8.65	0.76
- *Ave. ρ within SCEs*	*40%*	*0%*	*69%*	*58%*	*59%*	*47%*
(4) Total Value Gross of Subsidy(= (3)-(2)-(1))	270.26	286.97	295.61	274.88	274.82	266.65
Welfare (=(4) - (1+τ)*(3.2))	266.02	286.97	284.83	264.90	264.44	265.74
Return (Δwelfare/subsidy)	7.8%	NA	-24%	-10.1%	-15.0%	0.2%

Notes: This table reports four counterfactuals. Column 1 assumes that the government does not have the mild failure problem. It can tell the patent quality. Columns 2 and 3 double σ in the benchmark model (lower frictions in patent trade). In columns 4 and 5, the bureaucrats either only subsidizes in-house patents with $T = 0.1$ (column 2) or only subsidize in-house patents with a fixed budget (column 3). Column 6 assumes that the bureaucrats cannot tell patent quality, and then search for optimal subsidy intensities for in-house and purchased patents. p_H and p_L are in thousand RMBs. Other numbers except for percentage are in billion RMB.

Table 9: Extension - Possible Mislabeling of R&D Expenditure

	Laissez Faire	With subsidy	Count In-house patent
(1) Innovation Expenditure	6.15	6.82	7.09
(1.1) R&D cost for x_H	6.15	6.19	6.28
- SCEs with initial patents < 6	0.63	0.62	0.63
- SCEs with initial patents ≥ 6	1.70	1.64	1.70
- NCE in targeted industry	2.44	2.60	2.60
- Other NCE	1.38	1.33	1.35
(1.2) R&D cost for x_L	0.00	0.63	0.81
- SCEs with initial patents < 6	0.00	0.07	0.23
- SCEs with initial patents ≥ 6	0.00	0.08	0.02
- NCE in targeted industry	0.00	0.48	0.56
- Other NCE	0.00	0.00	0.00
Share of NCE pretend to be SCE	*0%*	*23%*	*22%*
(2) Value from Trade	0.00	0.00	0.00
- SCEs with initial patents < 6	-1.06	0.10	0.03
- SCEs with initial patents ≥ 6	-3.19	-3.64	-3.53
- NCE in targeted industry	0.19	-2.24	-1.68
- Other NCE	4.07	5.78	5.18
Price of high-quality patent p_H (in 1,000 RMBs)	350	330	330
Price of low-Quality patent p_H (in 1,000 RMBs)	0	4	0
(3) Revenue from Production and Subsidy			
(3.1) Output value excluding subsidy	271.76	271.29	275.22
- SCEs with initial patents < 6	14.97	10.72	10.36
- SCEs with initial patents ≥ 6	36.65	29.00	31.07
- NCE in targeted industry	77.17	86.61	89.62
- Other NCE	142.98	144.96	144.18
(3.2) Subsidy	0.00	20.93	20.17
- *Ave. ρ within SCEs*	*0%*	*59%*	*57%*
(4) Total Value Inclusive Subsidy =(3)+(2)-(1)	265.61	285.41	288.30
Welfare ((4) - (1+τ)*(3.2))	265.61	260.29	264.10
Return (Δwelfare/subsidy)	0%	-26.1%	-7%

Notes: This table describes an economy in which the NCEs may pretend to be SCEs. In the second column, the subsidy policy is the same as the previous baseline case. In the third column, the bureaucrats only subsidize in-house patents. p_H and p_L are in thousand RMBs. Other numbers except for percentage are in billion RMBs.

Figure 1: Data structure

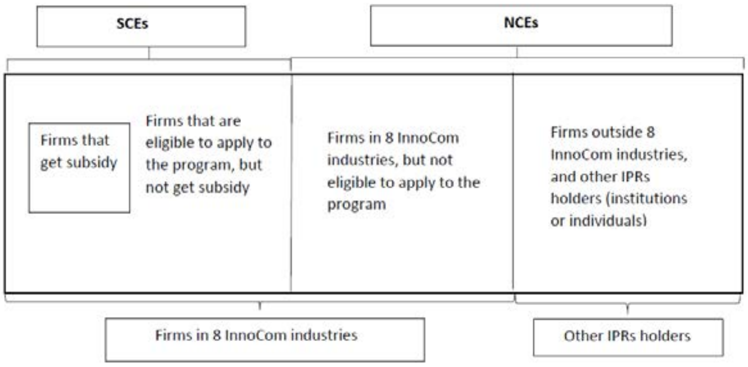
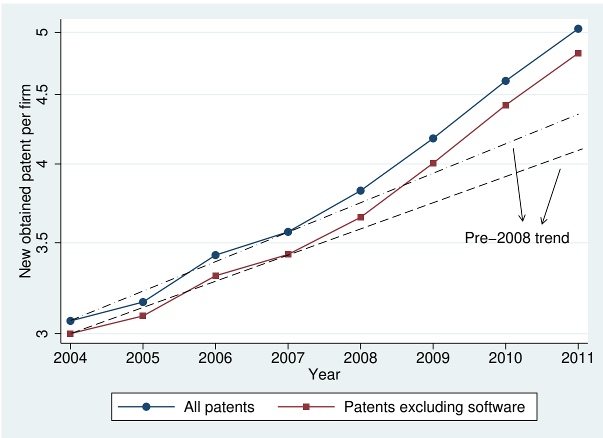
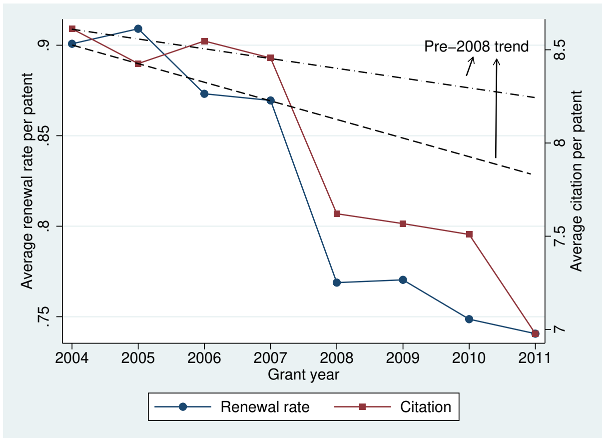


Figure 2: Patent quantity and patent quality



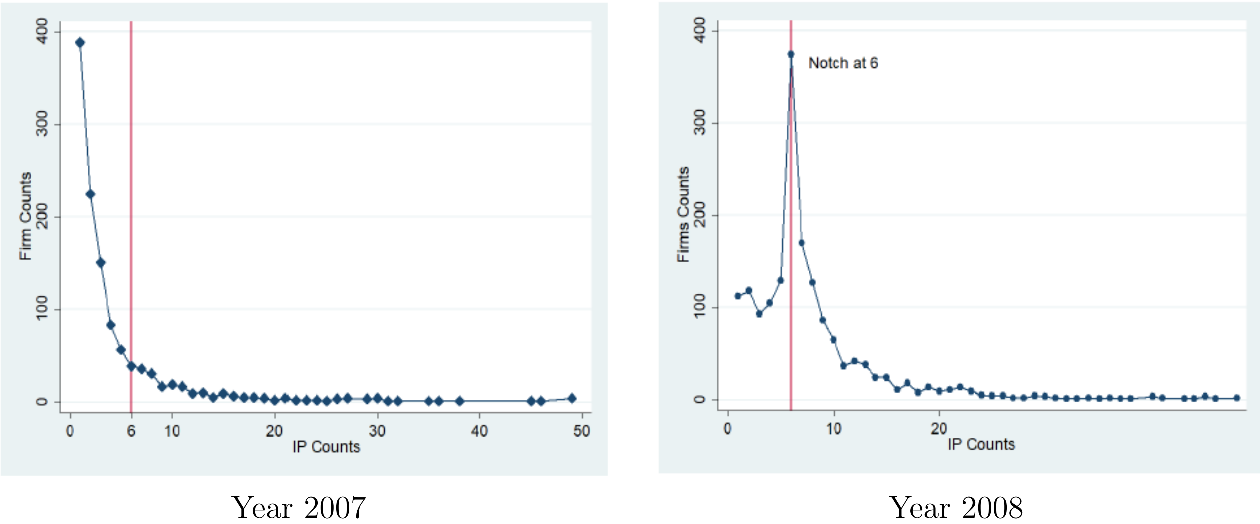
A: Patent quantity



B: Patent quality

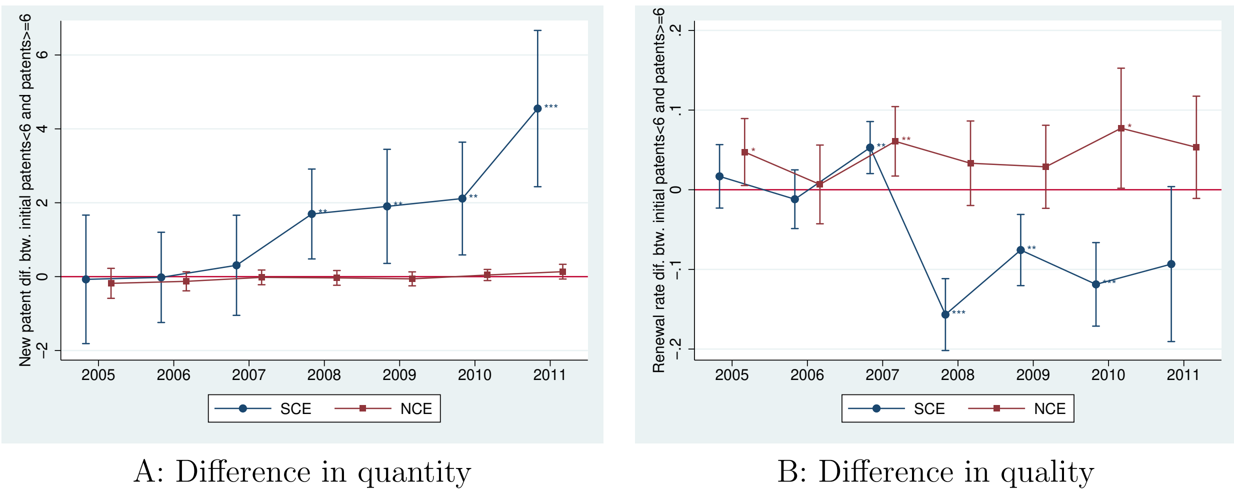
Notes: Figure A shows the number of new obtained patents per SCE from 2004-2011. Figure B shows the average 3-years renewal rates and citations of patents (invention and utility) granted within 2004-2011.

Figure 3: Patent count distributions of SCEs



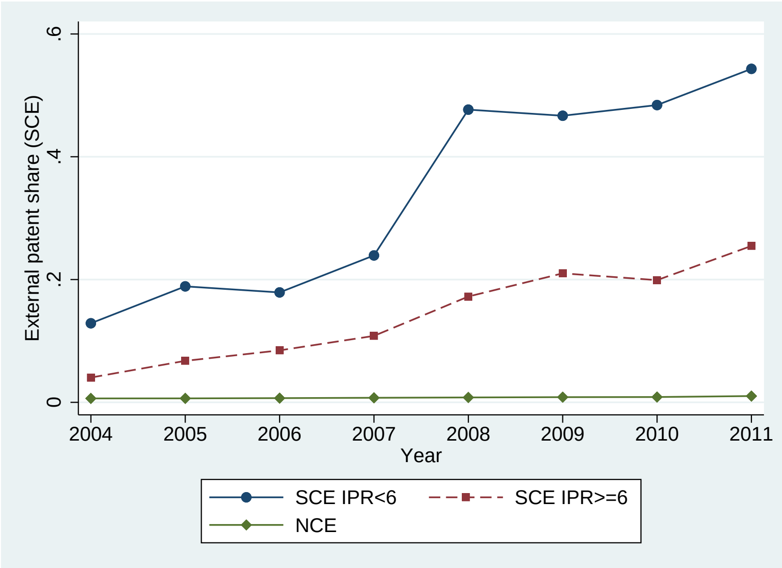
Notes: This figure shows the patents holding distributions of the subsidized firms in years 2007 and 2008.

Figure 4: Difference in innovation increase between SCEs with initially low and high patent counts



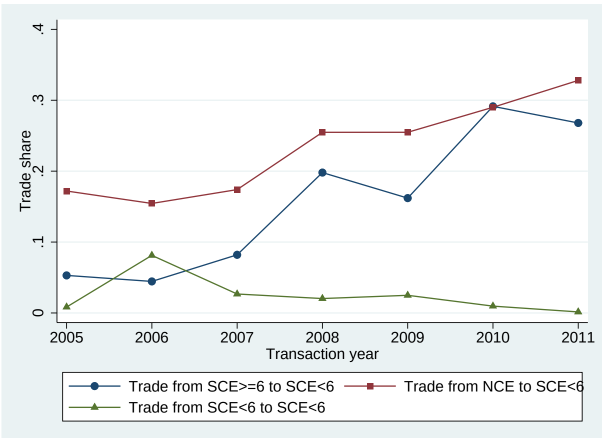
Notes: Figure A shows estimates of new patents difference between SCEs/NCEs with initial patents lower and higher than 6 (equation (2)) by years. Figure B shows estimates of renewal rate difference between SCEs/NCEs with initial patents lower and higher than 6 (equation (3)) by years. The circles indicate point estimates, and capped spikes represent 90% confidence intervals.

Figure 5: External patent share of SCEs and NCEs

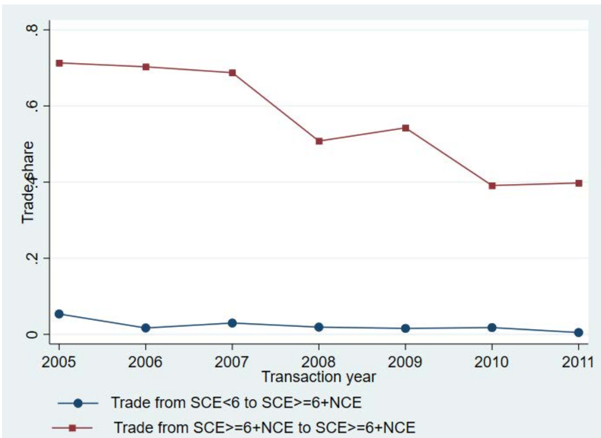


Notes: This figure shows the average external patent share of SCEs and NCEs.

Figure 6: Shares of patent sold to less innovative SCEs and other firm types



A: Shares sold to less innovative SCEs



B: Shares sold to other firms

Notes: Figure A shows the trade shares of patents sold to those SCEs with patent count < 6 from SCEs with initial patent count ≥ 6, NCEs, and SCEs with patent count < 6. Figure B shows the trade shares of patents sold to either the NCEs or the SCEs with patent count ≥ 6 from the SCEs with patent count $<$ or ≥ 6, respectively.

Figure 7: Timing of the model

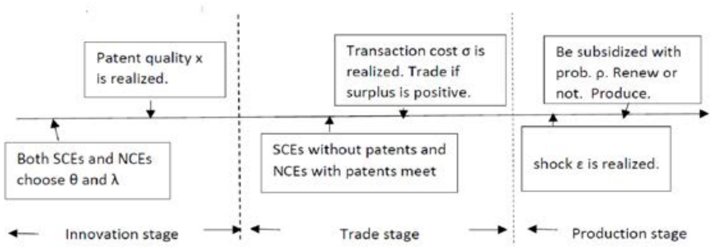
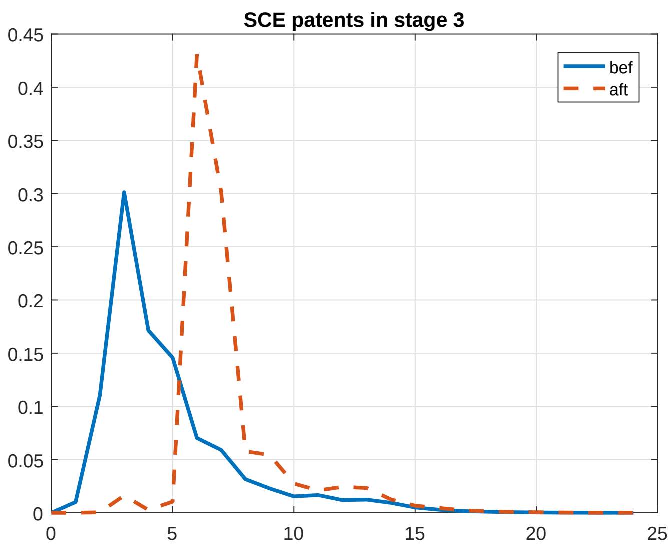
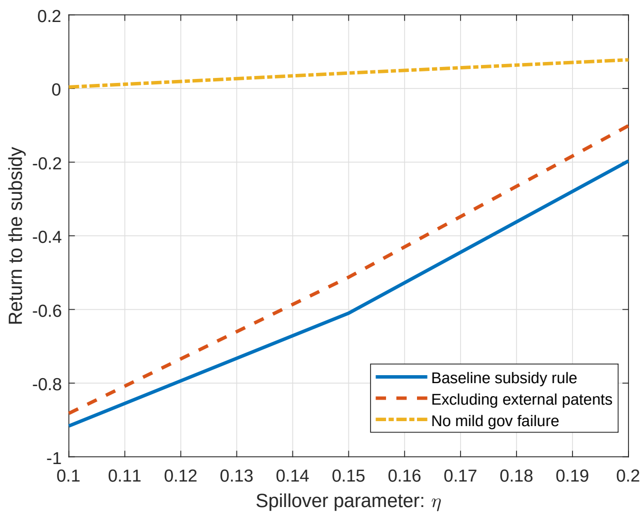


Figure 8: Model predicted distributions of patent counts for SCE before and after the subsidy



Notes: This figure shows the distribution of patent counts for SCEs before (solid line) and after (dashed line) the subsidy.

Figure 9: Return to the subsidy of InnoCom program under different technological spillover parameter η



Notes: This figure shows the return to the total subsidy when counting (solid line) or not counting (dashed line) external purchased patents.

Online Appendix (not for publication in print)

A When Innovation Outcome is Random

In this appendix, we reformulate the model so that the representative firm chooses the number of projects x and the probability to convert a project to a high-quality patent q, instead of choosing the numbers of high- and low-quality patents directly. We denote the expected numbers of high- and low-quality patents by θ_H and θ_L, respectively. Since $\theta_H = xq$ and $\theta_L = (1-x)q$, choosing x and q by the firm is equivalent to choosing θ_H and θ_L.

Define the numbers of high- and low-quality patents after the innovation stage by $\tilde{n}_H$ and $\tilde{n}_L$, respectively. Their expected values are $n_H + \theta_H$ and $n_L + \theta_L$, respectively. The firm's value at the beginning of the production stage is

$$E[V_3(i, z, \tilde{n}_H, \tilde{n}_L)] = \pi + E[\tilde{n}_H^\alpha]r(z) + E[\rho(i, \tilde{n}_H + \tilde{n}_L)]T\pi$$

where the expectation is over $\tilde{n}_H$ and $\tilde{n}_L$. Considering the first-order Taylor expansion around the expectations of $\tilde{n}_H$ and $\tilde{n}_L$, then this function can be written as

$$E[V_3(i, z, \tilde{n}_H, \tilde{n}_L)] \approx \pi + (n_H+\theta_H)^\alpha r(z) + \rho(i, n_H+\theta_H+n_L+\theta_L)T\pi = V_3(i, z, n_H+\theta_H, n_L+\theta_L)$$

This means that our benchmark model in the main text can be considered the first-order approximation of this model.

B Estimating Knowledge Spillover

Rather than relying on the literature for the value of knowledge spillover, We estimate it in this appendix. Denote N_{it} as the total patent count in the industry i cumulative up to year t, and r_{it} as the average renewal rate (survival rate of patents three years after the granted years) of the industry i. For every industry i in year t, we define a pair of within-sector knowledge capital, K_{it}^{within}, and knowledge capital from other sectors, K_{it}^{other}:

$$K_{it}^{within} = d_{ii}N_{it}r_{it},\ K_{it}^{other} = \sum_{j\neq i} d_{ij}N_{jt}r_{jt} \tag{13}$$

where d_{ij} is the knowledge distance between industry i and j, measured by the share of patent citations between industry i and j in all patent citations. The two functions imply that high-quality patents (high renewal rates) contribute more to the knowledge capital. The spillover from industry j to the industry i is also greater if the two industries cite each other more frequently.

We estimate the following equation

$$\ln \pi_{ft} = \eta_1 \ln K_{i_f,t}^{within} + \eta_2 \ln K_{i_f,t}^{other} + \mu_f + \mu_t + error \tag{14}$$

where π_{ft} is firm f's labor productivity (profit/workers) in year t, $K_{i_f,t}^{within}$ and $K_{i_f,t}^{other}$ are the two knowledge capitals facing firm f in industry i, respectively. μ_f and μ_t are the firm and time-fixed effects, respectively. η_1 and η_2 measure the spillovers from the within-sector knowledge accumulation, and other sectors' knowledge accumulation, respectively. As a firm's own innovation can change its productivity directly, we restrict our sample to firms without any patents. We also use only data during 2000-2007 to exclude the impact of the InnoCom program itself.[24]

As π_{ft} and industry-level innovations may respond to common shocks, we need to consider some instruments for $K_{i_f,t}^{within}$ and $K_{i_f,t}^{other}$, respectively. We assume that the changes in the patent count across industries in response to a subsidy program are proportional to the initial patent shares across industries. Furthermore, the subsidy does not change the productivity of no-patent firms directly. We construct our instruments as follows:

$$IV_{it}^{within} = d_{ii} \frac{Patent_{i0}}{TotPatent_0} S_t, \; IV_{it}^{other} = \sum_{j \neq i} d_{ij} \frac{Patent_{j0}}{TotPatent_0} S_t \tag{15}$$

where $Patent_{i0}$ and $TotPatent_0$ are the patent counts for industry i and the whole country, respectively, in 1998 (i.e., a decade before the subsidy program). S_t is the total subsidy to innovations.

The OLS estimates reported in the first column of Table B1 suggest that when the patent in a firm's own sector increases by 1%, the firm's labor productivity increases by 0.052%. But a weighted average of other sectors' patents increases by 1%, and the firm's labor productivity drops by 0.102%. In the second column, using the instrumented values

[24]We use the Annual Survey of Chinese Manufacture Enterprises database which covers firms' financial information up to 2009. We do not use the administrative tax records because it does not have data before 2007.

of the two knowledge capitals, we find that the spillover from own sector innovation is 0.01%, which is economically small. The spillover from other sectors' innovation is not statistically different from zero.[25]

As a sensitivity check, we refine the knowledge capital by simply adding up all patents without adjusting for renewal rates. The new estimates of η_1 and η_2, reported in Column 3, become smaller and neither is statistically significant.

To see if the knowledge spillover is stronger in the InnnoCom targeted industries than the economy at large, we interact a dummy for the InnoCom industries with the two knowledge capitals and report the results in Column 4. Neither coefficient on the two interaction terms is significant. We conclude therefore that the InnoCom industries are not special as far as the size of the spillover is concerned. We do the same regression using only the post-2008 data and find the same results (Column 5).

In Columns 6 and 7, we do similar regressions as in Columns 6 and 7 except that we do not adjust the patents by quality. (That is, knowledge capital is the simple sum of the patents regardless of the quality). In these cases, we see no positive and significant coefficients anymore. These results are consistent with the interpretation that low-quality patents do not contribute to productivity spillover.

In sum, we find a positive but small knowledge spillover from within-sector innovations by other firms, but no significant spillover from innovations in other sectors. Note that the estimated size of spillover in our data is smaller than those in Lucking et al. (2019).

[25] In the first stage, we find the F statistics is over 2,000.

Table B1: Knowledge Externality of Patents

	(1) OLS	(2) IV	(3) IV raw K	(4) IV InnoCom Ind.	(5) IV InnoCom Ind. Post-2008	(6) IV InnoCom Ind. raw K	(7) IV InnoCom Ind. raw K / Post-2008
$\ln K^{within}$	0.052***	0.010*	0.009	0.011*	0.015*	0.01	0.011
	-0.012	-0.006	-0.008	-0.007	-0.008	-0.008	-0.013
$\ln K^{other}$	-0.102***	-0.012	-0.013	-0.011	-0.024	-0.013	-0.009
	-0.025	-0.016	-0.017	-0.019	-0.017	-0.019	-0.026
$\ln K^{within} \times D_{InnoCom}$				-0.005	-0.013	-0.004	-0.015*
				-0.008	-0.009	-0.008	-0.009
$\ln K^{other} \times D_{InnoCom}$				0.008	0.003	0.007	0.004
				-0.009	-0.009	-0.009	-0.009
Firm FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y
IV		Y	Y	Y	Y	Y	Y
K: Renewal rate adj.	Y	Y		Y	Y		
Sample after 2008					Y		Y
KP Wald F stat		2175.7	3017.11	5929.45	1500	4594.17	1400.04
Obs.	1,120,443	1,120,443	1,120,443	1,120,443	373,074	1,120,443	373,074
Adj. R2	0.65	0.65	0.65	0.65	0.67	0.65	0.67

Notes: This table reports the estimates of knowledge spillover within and across sectors (equation (14)). The dependent variable is the log of labor productivity (profit/worker) of a firm without a patent. The $D_{InnoCom} = 1$ if the firm is in an InnoCom industry and 0 otherwise. In column (3), K is not adjusted by the renewal rate. The IV of columns (2) to (7) is defined in equation (15). “KP Wald F stat” refers to Kleibergen-Paap Wald F-statistic in the first stage estimation. Standard errors are reported in parentheses and clustered at the firm-year level. *** p<0.01, ** p<0.05, * p<0.1.