Bureaucratic Bean Counting and Patent Subsidies: A Welfare Evaluation of China's Pro-innovation Program

Shang-Jin Wei [*], Jianhuan Xu [†], Ge Yin [‡], and Xiaobo Zhang [§]

January 13, 2021

Abstract

This paper quantifies the welfare effect of the patent subsidy policy (the InnoCom program) in China. The policy encourages firms to hold more patents, but disregard the quality of patents. We provide evidence that the subsidy decreases the value of patents either through declining the quality of new granted patents or through the inefficient trade of patents. We build a structural equilibrium model, which allows three channels that the subsidy policy can change the aggregate welfare: (1) Distorting the subsidized firm's innovations by increasing low quality new patents; (2) Reallocating patents to inefficient firms for being qualified to be subsidized; (3) Imposing a positive demand shock on the patent trade market and encouraging more low quality new patents. We find that the aggregate welfare declines substantially and the third channel is the most important one, suggesting that it is crucial to jointly consider the patent trade and the entry of new patents.

[*] Columbia University, FISF, NBER, and CEPR

[†] Singapore Management University

[‡] Peking University

[§] Peking University and IFPRI

1 Introduction

Firms may under-invest in innovation when there is no government intervention. Given that intellectual property rights (IPRs), especially the patents, have long been considered as one of the most important measures for evaluating innovation, lots of governments have put financial supports on fostering patent initiatives and exploring patented technologies.[1] For instance, more than 10 countries, including UK and France, have patent box policies to give a very low corporate tax to the patent holders. However, little is known about the actual welfare impact of these government-backed support policies.

Among these countries, China's IPRs subsidy policy is a dazzling example. For more than a decade, Chinese government has set up a variety of incentives, called the InnoCom Program, to encourage firms to create and utilise more IPRs. The consequence of the Chinese government IPRs subsidy policy is deeply impressive. From 1995 to 2014, the number of patent filings has increased with an annual speed around 20% (Wei et al. (2017)). However, without taking into account the quality of patents, the impressive patent growth rate does not tell us the actual innovation in China directly.[2] This paper aims to quantify the welfare impact of the Chinese innovation subsidy policy: the InnoCom program.

There are two unique features of the InnoCom program. First, the program chooses firms to subsidize based on the number of IPRs, disregard the quality of them. As long as the firm has 6 or more IPRs, it will get a grade "A" in the evaluation. Hence the firm wants to increase its number of IPRs up to 6, but cares less on the value of IPRs. We utilize this policy feature to identify the causal effect of the the InnoCom program on the value of innovations: the firm's innovation response to the policy may be significantly different around 6 IPRs.

[1] The IPRs include patents and other non-patent intellectual property rights, such as software copyrights. In this paper, we will mainly focus on patents.

[2] Some anecdotal evidence suggests that the quality of patents in China may decrease. For instance, there is only 26% of Chinese patent applications are granted, which stands in sharp contrast to the 51% of US patent granted rate. This may suggest the low quality of patents in China.

Second, both the self-developed IPRs and the external acquired IPRs are counted in the same way by the InnoCom program.[3] Firms can buy patents in order to get the subsidy. So the subsidy policy not only changes the innovation of subsidized firms directly, but also indirectly change the innovation of non-subsidized firms through the patent trade market. Although only a relative small share of firms get the subsidy directly (7% in Beijing), the policy may have a large aggregate impact by introducing a positive demand shock on the patent trade market.

The paper has two parts. We first provide some reduced-form evidence on the welfare impact of the subsidy. We find evidence that the policy decreases the innovation efficiency for both subsidized firms and non-subsidized firms. For subsidized firms, we find that they increase their patents either by self-development or purchasing from others.[4] Especially, it their IPRs are lower than 6, the increase of patents is much more pronounced. Meanwhile, recognizing that renewal rate positively correlates with the value of patents, we find that the renewal rate declines after the subsidy policy, which suggests that the value of patents declines.[5] Interestingly, the decline of the renewal rate is only pronounced among firms whose IPRs are lower than 6, which have the strongest incentives to increase the number of patents, disregard the quality of patents.

For non-subsidized firms, the subsidy introduces a higher demand for their patents. We find that if they sell more patents in the previous year, they grant more low quality patterns. To identify the causal relation between the trade of patents and the increase of low quality

[3]There are two ways for the firms to acquire an IPR externally: purchasing it or signing an exclusive usage license. We do not distinguish these two in the paper.

[4]We do not have good information on the IPRs, except for patents. So we focus on patents in this paper.

[5]Using the renewal rate to measure the value of patents is first explored by Pakes (1986). The idea is that if the patent has a low value, it is less likely to be renewed. There is a large literature follows this idea (Cornelli and Schankerman (1999), Schankerman (1998), Lanjouw (1998), Schankerman (1998), Deng (2007) and Bessen (2008)).

patents, we follow a Bartik et al. (1991) instrument strategy.[6] Firms with fewer than 6 IPRs have strong incentives to purchase patents. Before the subsidy policy, the share of patents sold to firms with fewer than 6 IPRs is different across non-subsidized firms. It gives an exogenous exposure for different firms to the increase of the patent trade. We instrument the sale of patents from non-subsidized firms by using the initial trade composition and the aggregate patents trade within subsidized firms. We find the pattern is even more pronounced.

In the second part of the paper, we quantify the welfare impact of the InnoCom program by developing an equilibrium model. The model highlights three channels that the policy can affect the welfare: (1) The direct effect. Those to-be-subsidized firms' may want to grant more patents to get subsidies, but do not care the quality of patents. (2) The reallocation effect. A patent may be sold to an inefficient firm solely because it has the chance to get the subsidy. (3) The equilibrium effect. Firms which have no chance to get the subsidy may grant more low quality patents because the increase of the patents demand.

After fitting the model to the data, we find that the InnoCom program decreases each patent's value by 321 RMB on average. The aggregate revenue that the patents can create declines by 0.53 billion RMB per year, which is about 42% of the total patents' revenue before the subsidy. Our model also suggests that the most important channel to determine the aggregate welfare loss of the policy is the trade effect. It alone can explain about about half of the aggregate welfare decline.

We are not the first one to study the patent subsidy program in China (Li (2012), Dang and Motohashi (2015) and Zhao et al. (2019)). However, most scholars in this literature try to build the causal connection between the surge of patents in China or the quality change of new patents and the subsidy. Our paper goes beyond the literature by adding a new key

[6]See Goldsmith-Pinkham et al. (2018) for a detailed explanation on the Bartik IV.

element: the patent trade market. Importantly, we find that most welfare impact of the subsidy comes from this channel.

Our paper also relates to several other streams of literature. First, there is a large literature that estimates the private returns of patents using the renewal information starting from Pakes (1986). However, most papers in this field focus on the gains of patent protection (such as increasing the R&D incentives) and do not consider the trade of the patents.

Some other papers study the patent trade market, which try to identify the gains from the patent trade or highlight frictions on this market.[7] However, most papers do not quantitatively evaluate the patent trade market efficiency, with two exceptions: Serrano (2018) and Akcigit et al. (2016). Serrano (2018) use the patent renewal rate information to estimate the value of patent trade but does not endogenize the entry of new patents. Akcigit et al. (2016) builds an endogenous growth model and estimate how the search friction on the patent trade market can affect the growth. Our paper distinguishes from their works by studying a different source of friction: the patent subsidy.

Finally, the paper relates to the empirical literature on the industrial policy, especially the R&D subsidy.[8] Bloom et al. (2002) and Wilson (2009) study the impact of the tax credit on the level of R&D investment in US. A special feature of the Chinese patent subsidy is that it not only changes the firm's self-developed innovation directly, but also change the innovation through the patent trade market.

The paper is organized as follows: Section 2 introduces the details of the InnoCom

[7]For instance, Gans and Stern (2000) and Serrano (2010) point out the gain of patent trade comes from different firms have different abilities to commercialise the patents; Galasso et al. (2013) argue that the patent trade can help resolve disputes without resorting to the courts. In terms of the frictions of the patent trade market, Shapiro (2010) and Lemley and Shapiro (2005) argue that the the patents may be reallocated to entities that have the capacity to exploit the patents to extract excessive royalty fees by holding up rivals. Akcigit et al. (2016) highlight the trade market has a large search friction.

[8]Examples of research on other industrial policies include export subsidies (Das et al. (2007)), subsidies to places with poor natural conditions(Neumark and Simpson (2015)), environmental subsidies (Yi et al. (2019)). Some papers study the industrial policies in China, such as Liu (2019) and Barwick et al. (2019).

program and Section 3 explains the data we use. The reduced form evidence on the welfare impact of the subsidy policy is provided in Section 4. Some empirical patterns of the patent innovation and trade will be shown as well, which motivate the assumptions of our structural model which is laid out in Section 5. We then fit the model to the data in Section 6 and the counterfactual analysis are shown in Section 7. Finally, Section 8 concludes the paper.

2 The Chinese InnoCom Program

The Chinese InnoCom program, which was introduced since 2008, is intended to encourage the firm's innovation. The program targets qualified high-tech firms and awards them financial support and favorable policies. The most important benefit is a 15% corporate income tax cut, which is equivalent to a 40% total corporate tax reduction.

The selection of the qualified high-tech firms is determined by the provincial government. In Beijing, the selection process has two steps. In the first stage, the firms from 8 industries, which satisfy the pre-required conditions, apply to the local government.[9] In the second stage, the qualified firms will be evaluated by the government from four categories: the number of IPRs, the ability to manage R&D, science and technology commercialization ability, and the firm growth potential. The full marks in these 4 categories are 30, 30, 20 and 20 respectively. If the firm scores 70 points out of 100, it will be awarded the title of *InnoCom firm*, and the corporate income tax rate is cut. These subsidized firms are reviewed every three years to determine whether they can stay in the program.[10]

[9]These 8 industries are the pharmaceutical manufacture (CSIC 27), the special equipment manufacture (CSIC 36), the transportation equipment manufacture (CSIC 37), the communication equipment & computer manufacture (CSIC 40), the precision instrument manufacture (CSIC 41), computer service (CSIC 61), software service (CSIC 62) and environmental protection industry (CSIC 80). We call them as InnoCom industries. The pre-requisites are based on the R&D-to-sales ratio, the sales of high-tech products, the share of R&D workers, and the number of IPRs.

[10]The details of grading criteria of each category are provided in Appendix A.

A noteworthy point is the grading scheme of the IPRs. The firm gets an "A" if they have 6 or more IPRs. Here we emphasize two important points. First, the subsidy program only counts the number of IPRs but does not take into account the quality of the IPRs. This may distort the people's innovation behaviors. Second, firms do not get any extra points if they have more than 6 IPRs. We argue that these policy properties drive a substantial welfare loss.

The InnoCom program is one of the most aggressive innovation subsidy in China. In the first column of Table 1, we list the number of new InnoCom firms in Beijing from 2008 to 2012. The programs subsidized more and more firms per year, from 2,608 firms in 2008 to 7,920 firms in 2012. On average, the Beijing government subsidized 5,860 firms per year. In the second column, we list the number of patent holders in Beijing, defined as Beijing patent holders (including firms, individuals or institutions) which are not covered by the InnoCom program. As we can see, there are about 84,091 patent holders on average. The InnoCom program subsidized about 7% (5,860/84,091) of patent holders annually in Beijing. In the last column, we report the annual subsidy per InnoCom firm, about 0.618 million RMB.[11] Hence the total subsidy is about 3.4 billion RMB per year (0.61*5,860). Meanwhile, considering these subsidized firms could purchase patents from other patent holders, the number of firms that are affected by the subsidy could be much larger.

In this paper, we quantify the Beijing subsidy policy. However, we want to point out that the policy is used nation-wide. According the Xinhua News Agency, up to 2019, there are over 181,000 InnoCom firms in China. To give a sense on the magnitude, the number of above-scale firms in China (annual sales above 20 million RMB) is about 300,000, among

[11]We infer the subsidy as follows. We first find InnoCom firms in the Chinese State Administration of Tax Data, which covers the detailed tax information at the firm level, by matching firms' names. We can match 2,272 subsidized firms in the tax data per year on average from 2009 to 2011. But we do not find any InnoCom firms in 2008 and 2012 in the tax data. The InnoCom subsidy is defined as 15% of rebate of the corporate tax.

which about 47,000 are from the 8 InnoCom industries.[12]

3 Data

We mainly use three data sets. First, we have the detail information of those firms who were subsidized by the program in 2008. The data is collected by the Beijing Municipal Science & Technology Commission. It includes the grading information of the four categories such as number of IPRs, number of research workers and other related financial information in 2008. Panel A of Table 2 summarizes the key variables of the 2008 subsidized firms. From the first four rows, we can see that there are 2,608 firms got subsidized in Beijing. Most firms are large in terms of sales, assets (above 14 million RMB and 18 million RMB on average) and number of workers (above 160 employees on average). They are research intensive firms. The share of research staffs is about 44%. In the next three rows of Panel A, we report the summary statistics of the number of IPRs and patents. IPRs include not only patents, but also copyrights, logos and etc. We can see that on average the firms have 8.8 IPRs, among which 3.5 are patents and 5.3 are non-patent IPRs. Within the non-patent IPRs, the software copyright is the most important one. On average, a firm holds 5.1 software copyrights, which is more than 96% of the non-patent IPRs. In the last row, we report their evaluation scores from the government. Since all of them are subsidized, their scores are all above 70. The average score is about 81 and the standard deviation is about 7.2.

We also have the name list of all subsidized firms from 2009-2015 in Beijing, but lacking other detailed information.

The second and the third data come from the China's State IPRs Office. First, it provides the patent assignment information from 2001-2015, which records the patent grant and

[12]Using the Tax Administration data, we find the national average annual subsidy per firm is about 1.28 million RMB, which is twice of the value in Beijing.

transaction information. We define that a patent is re-allocated to the other firm either the patent is sold or it is used under an exclusive license.

Although we have lots of information about patents, we do not have very good information about other IPRs. The State IPRs Office only provides us the software copyrights assignment information from 2007-2008. Using this data, we can only get the number of software of each firm in these two years. Noticing that the software is the most important non-patent IPRs (Table 2), we use the number of software to measure the number of non-patent IPRs of each firm.

To quantify the welfare effect of the Beijing government's subsidy policy, we restrict our attention to those patent holders that located in Beijing since they are mostly affected by the subsidy program.[13]

We study the subsidy policy from three aspects: the direct subsidized firms, the re-allocation on the patent trade market and other non-subsidized firms. First, we focus on subsidized firms covered by the InnoCom program in 2008. We utilize the rich firm level information of the 2008 subsidized firms and merge it with the patent assignment data as well as the software assignment data. Hence we can construct the patent holding portfolio of each 2008 subsidized firm over years and the software copyrights from 2007-2008. We use firm names to merge several data sets. Overall, we can find 1,776 firms out of the 2,608 subsidized firms show up in the IPRs assignment data.[14] Panel B of Table 2 reports the summary statistics of the matched sample. Comparing with Panel A, the matched firms are larger than other subsidized firms in terms of sales, assets or the number of workers, but the

[13] A Beijing firm can buy patents from people of other cities. Hence patent holders outside Beijing may also be affected by the subsidy program. However, we find that most patent trade (82% of the trade) are within the same city.

[14] Some subsidized firms only have non-patent IPRs in 2007. The matching rate between these firms and the software assignment data is only about 36%. It may be because that their IPRs are not software. The matching rate between subsidized firms with patents and the patent assignment data is about 94%, similar with the matching result in Hu and Jefferson (2009). The overall matching rate is 1,776/2,608=68%. The details of the matching procedure are provided in Appendix B.

share of research staff is lower. However, none of these differences are significant. Moreover, we can see that their scores are very close to the scores of other subsidized firms. So we conclude that the matched sample is not significant different away from other subsidized firms.

When we study the patent trade market and other non-subsidized firms' innovation behaviors, we use the patent assignment data from 2001-2015 and merge it with the name list of subsidized firms from 2008-2015. So we can construct the patent portfolio for each firm as well as the patent re-allocation from 2001 to 2015.

4 Impacts of the InnoCom Program: Reduced-form Evidence

In this section, we provide reduced-form evidence on the impacts of the subsidy from three aspects. First, we focus on those subsidized firms. Second, we study the patent trade market. Finally, we check the the policy's impact on granting new patents of non-subsidized firms .

4.1 The impact of the subsidy on subsidized firms

Since we have detailed information for firms subsidized in 2008, we only study this group of subsidized firms in this subsection. We focus on their patent portfolio changes from 2001 to 2010 to isolate the one time subsidy effect, since the 2008 subsidized firms were reviewed in 2011. We show that subsidized firms increase their patents, especially low value patents after the subsidy policy.

Subsidized firms get more IPRs up to 6

As discussed in the previous section, the firm gets no more extra point when its IPRs are above 6. Thus we expect to see that subsidized firms may get IPRs up to 6. We sum the patent and the software copyrights to get the number of IPRs by firm in 2007 and 2008. In Figure 1, we plot the distribution of the IPRs counts. The left graph corresponds to the IPR distribution in year 2008. There is a significant amount of subsidized firms which have just 6 IPRs. In the right, we plot the IPR distribution in year 2007. The distribution follows a smooth decline pattern, without a mass point around 6. Comparing these two graphs, we see that firms strategically respond to the policy by adjusting their IPRs holdings up to 6.[15]

We then switch to patents. In Figure 2, we plot the average of new patents the firm gets per year from 2002 to 2010 in the solid line.[16] Overall, firms get more and more new patents per year, from close to 0 new patent in 2002 to about 2.3 patents in 2010. However, there is a peak in the year 2008. The firm gets 2.4 new patents in 2008, jumping from 1 new patent in 2007. Then after 2008, the number of new patents declines over time. This confirms that firms try to get more patents in response to the InnoCom program.

To address the endogeneity concern, we adopt the following strategy. We first compute the number of IPRs (the sum of patent and the copyrights of software) in year 2007. From Figure 1, we see that firms whose IPRs are lower than 6 in 2007 have stronger incentives to increase their patents. Thus firms with lower than 6 IPRs should increase patents more in 2008.

[15] The InnoCom program was formally announced in April of 2008. Some firms may know the program before the announcement and started to adjust their patents before 2008. However, none of firms know the details of the evaluation criteria and the scale of the subsidy. Given the distribution in 2007 does not have a mass point around 6 IPRs, we choose year 2008 as the time when firms change their behaviors in the following analysis.

[16] To compute the new patents, we need to compare two years' patent holding. So we can only start from year 2002.

Formally, we estimate the following equation:

$$N_{it} = \alpha_0(IPR_{2007} < 6) \times T_{t=2008} + T_{t=2008} + X_{it} + \mu_i + \epsilon_{it} \tag{1}$$

where N_{it} is the number of new patents (either from self-innovation or purchased from others) of firm i in year t. $(IPR_{2007} < 6)$ is an indicator function, which equal to 1 if the number of IPRs in 2007 is smaller than 6, and 0 otherwise. $T_{t=2008}$ is another indicator function, which equal to 1 if the year is 2008 and 0 otherwise. X_{it} is the control at the firm-year level, such as the last year's number of patents. μ_i and μ_t are the year and firm fixed effects. ϵ_{it} is the error term. α_0 measures the difference of the new patents between $(IPR_{2007} < 6)$ firms and $(IPR_{2007} \geq 6)$ firms.

Table 3 shows the results. Column 1 regresses N_{it} on the year indicator T and controls the firm fixed effect. A positive significant coefficient means that in 2008, the firm gets 0.655 more patents than other years. In Column 2, we regress N_{it} on the interaction term between $(IPR_{2007} < 6)$ and T. A positive coefficient before the interaction term suggests that if the number of IPRs in 2007 is less than 6, the number of new patents in 2008 is 0.474 more than other firms. In the third column, we have a more aggressive specification by controlling the year fixed effect. In this case, the dummy $T_{t=2008}$ is absorbed by the year fixed effect. While the interaction between $(IPR_{2007} < 6)$ and T is still significantly positive. The fourth column shows a placebo test. We define a new year indicator $\tilde{T} = 1$ if the year is 2007 and 0 otherwise. Estimating a similar equation with Column 2, we find that for those firms whose IPRs are lower than 6, their new patents are 1.487 lower than other firms, the opposite pattern comparing to that in 2008. This may be driven by that the low IPRs firms' innovation abilities are lower than those of high IPRs firms. Without the incentives to get the subsidy, low IPRs firms have fewer new patents than high IPRs firms.

In all specifications, we control the number of patents in the last year and the results are quite consistent in all specifications. On average, the last period patent and the patent increase are positively correlated. This may be driven by the firm's time-varying research abilities.

Within firms whose IPRs are lower than 6, if their IPRs are far below 6, their incentives to increase the patents should be stronger. Thus we should see that if the 2007 IPRs are higher (but still lower than 6), their new patents may be smaller. However, for firms whose IPRs are greater than 6, since there are no extra benefit of getting more patents, there should be no correlation between the number of initial IPRs and the number of new patents. We augment the estimation in Column 2 of Table 3 by estimating α_0 for different IPRs. Figure 3 plots the coefficients. The circle and triangle denote the point estimates of the coefficient when $IPR < 6$ or $IPR \geq 6$ respectively. We also show the 95% intervals on the graph. The x-line is the distance away from 6 IPRs. For instance, if x-line's value is 2 or 3, it means that the number of IPRs is either 3 to 4 (circle) or 8 to 9 (triangle). First consider $IPR < 6$ (circles). If the number of IPRs is lower than 4, firms would significantly increase their patents about 0.5 to 0.8. The lower initial IPRs are, the more patents would be obtained in 2008. When the number of IPRs is close to 6, firms do not change their patents significantly. However, when $IPR \geq 6$, all coefficients are not significantly away from 0.

The value of patents decline for the subsidized firms

We now switch to the impact of the policy on the efficiency of the subsidized firms. Specifically, we want to compare the value of patents obtained in 2008 with the value of patents obtained in other years. How to measure the value of patents? Following Pakes (1986) and Serrano (2018), we measure the value of patents by the renewal rate. The idea is that given there is a cost to maintain the patent, the low value patent is less likely to be renewed.

We consider three empirical specifications. First, we test whether patents obtained in 2008 are less likely to be renewed comparing to patents obtained in other years.

$$V_{ijt} = \alpha_0(T_{ij} = 2008) + Z_{ij} + X_{jt} + \mu_{it} + \epsilon_{ijt} \tag{2}$$

where i and j denote the owner of the patent (the firm) and the patent respectively. t denotes the year. V_{ijt} is an indicator that equals to 1 if the patent j is renewed in year t and 0 otherwise. T_{ij} is the year that the firm i gets the patent j. For patents that are developed by the firm, T_{ij} is the granted year, and for patents that are purchased from others, T_{ij} is year of the trade. $(T_{ij} = 2008)$ is an indicator function, which equals to 1 if T_{ij} is 2008 (the firm gets the patent in 2008) and 0 otherwise. The coefficient α_0 measures the renewal rate difference between patents that are obtained in 2008 and other years. Z_{ij} denotes whether the patent j is acquired (equals to 1) or self-developed (equals to 0). This control is motivated by the concern that patents selected to trade may be different from the self-developed patents. X_{jt} are observed characteristics of the patent j, including patent's age and the patent's IPC (International Patent Classification Code)-year fixed effect. μ_{it} denotes the firm-year fixed effect which tries to control the patent owner's productivity shock.

The result is shown in the first column of Table 4. First, $\alpha_0 = -0.9\%$ and is significant at 10% level, suggesting that patents which are obtained in year 2008 is less likely to be renewed. In terms of the magnitude, for patents obtained in 2008, the renewal rates are about 0.9% (about 1% of the average value) lower than those of patents obtained in other years. Second, the negative coefficient before Z_{ij} suggests that if patents are acquired from others, they are less likely to be renewed. However, the difference is not significant between external acquired patents and self-developed patents. Finally, old patents are less likely to be renewed.

Then we use a similar strategy as before to identify the causal effect of the policy on the renewal rates. For firms which have strong incentives to increase their patents (IPRs lower than 6), they may care less on patent qualities. For these firms, new patents obtained in 2008 may be less likely to be renewed. We estimate the equation (2) by adding an interaction term $(IPR_{2007} < 6) \times (T_{ij} = 2008)$, which equals to 1 if the patent j is obtained by the firm i in year 2008 and the firm i has fewer than 6 IPRs in 2007, and 0 otherwise. Column 2 of Table 4 shows the estimation result. Consistent with our expectation, the coefficient before $(IPR_{2007} < 6) \times (T_{ij} = 2008)$ is -0.018 , meaning that the renewal rates of low IPR firms' new patents in 2008 are 1.8% lower than that of high IPR firms. The coefficients before the Z_{ij} and the patent age are quite similar to those in Column 1.

In the third column of Table 4, we further control the year that firms get the patents (called the transaction year fixed effect). By doing so, we allow the patents obtained in different years have different renewal rates. However, coefficient before the interaction term does not change.

In the fourth column of Table 4, we show a placebo test. We change the indicator function $(T_{ij} = 2008)$ in the equation (2) to $(T_{ij} = 2007)$. We find that there is no significant different between the renewal rates of firms whose IPRs are less than 6 and those of other firms.

Finally, we test whether there is any difference in renewal rate within the firms with lower than 6 IPRs. We estimate a similar equation with Column 2, but add two triple interaction terms $IPR_{2007} \times (IPR_{2007} < 6) \times (T_{ij} = 2008)$ and $IPR_{2007} \times (IPR_{2007} \geq 6) \times (T_{ij} = 2008)$. Their coefficients measure the change of renewal rates when IPR_{2007} increases by 1, conditional on the $IPR_{2007} < 6$ and $IPR_{2007} \geq 6$ respectively. Since firms whose IPRs are far below 6 have stronger incentives to increase their patents, their new patents have lower qualities and lower renewal rates. Hence we see a positive coefficient before $IPR_{2007} \times (IPR_{2007} < 6) \times (T_{ij} = 2008)$. which means that when the number of IPRs increases, the renewal rates of 2008 new patents will increase as well. However, because there are no extra incentives for firms to increase their patents when the IPR is greater than 6, there is no significant correlation between the number of IPRs and the renewal rates of 2008 new patents among firms with more than 6 IPRs.

4.2 The impact of the subsidy on the patent trade

The subsidy policy may induce inefficient patent trade. For instance, the patent may be sold to an unproductive firm solely because the firm has a chance to get the subsidy. In this subsection, we show that (a) the number of patents trade increases since 2008; (b) the quality of newly traded patents since 2008 is lower than that in earlier years. We utilize all the patent assignment data from 2001 to 2015.

We categorize the trade into four groups according to the number of patents of the buyer and seller one year before the transaction year. Specifically, in the first group of patent transactions, the buyers have fewer than 6 IPRs before the transaction and the sellers have more than 6 IPRs. In the second group, both buyers and sellers have fewer than 6 IPRs. In the third group, both buyers and sellers have more than 6 IPRs. And in the last group, the buyers have more than 6 IPRs and sellers have fewer than 6 IPRs.[17] Under the subsidy policy, we expect the first group of transactions would increase the most since they have the most strong incentives to trade, while the number of transactions in the fourth group would increase the least.

Figure 4 confirms our conjecture. We plot the increase of the traded patents after 2008, comparing to that before 2008. Each point is the change of the transaction counts of a group.[18] As we can see, the transaction of the first group (firms with more than 6 IPRs sell

[17]Unluckily, we only have the software counts in 2007 and 2008. So we use the average of software in these two years + the number of patents to construct the number of IPRs at the firm-year level.

[18]We define a dummy "post"=1 if the transaction year is after 2008, and 0 otherwise. We then regress

to firms with lower than 6 IPRs) increases the most. The trade between a buyer-seller pair increases about 0.6 after 2008. Meanwhile, the increase of the transaction from the fourth group (firms with fewer than 6 IPRs sell to firms with more than 6 IPRs) is only around 0.17. The increase of trade of other two groups (trade between firms with more than 6 IPRs sell, or within firms with fewer than 6 IPRs) is in the middle, but not significantly different from each other.

We then focus on the quality of patents that are traded. We measure the quality of the patents using the renewal rate 3 years after the transaction.[19] Our conjecture is that the quality of the patents that are purchased by firms with fewer than 6 IPRs would decline the most after 2008, since these buyers have the strongest incentives to increase the number of patents, and care less about the quality of patents.

Figure 5 shows the result. We plot the renewal rate of the traded patents after 2008, comparing to that before 2008. Each point is the change of the average renewal rate within a group.[20] We can see that renewal rates of all trade decline after 2008. For patents that are sold to firms with lower than 6 patents, the renewal rate would decline most by 8%-10%. However, the decline is not significant when patents are sold to firms with more than 6 IPRs. Meanwhile, within the first two groups, if the patent is sold from a firm with more than 6 IPRs to a firm with lower than 6 IPRs (group 1), the decline of the renewal rate is more pronounced than that when the patent is sold from a firm with fewer than 6 IPRs to a firm with fewer than 6 IPRs (group 2). This may be because that the incentive to sell patents is stronger within firms with more than 6 IPRs.

---

the number of traded patents on the interaction term of the four group dummies and the post dummy, controlling the buyer and seller fixed effects.

[19]We use renewal rate 3 years after the transaction because the subsidized firms are under the review every 3 years.

[20]We regress the renewal rate 3 years after the transaction on the interaction term of the four group dummies and the post dummy, controlling the buyer and seller fixed effects.

4.3 New granted patents of non-subsidized firms

Facing a high demand of patents, firms which do not have chances to be subsidized may want to grant more patents. In this subsection, we show that (a) the subsidy policy increases the grant of new patents even within firms which never have chances to get subsidy; (b) the new granted patents are low quality patents.

We first look at the number of new granted patents of those never subsidized firms. When the policy increases the trade, these non-subsidized firms have stronger incentives to increase their patents. Formally, we estimate the following equation,

$$N_{it} = \gamma_0 Trade_{it-1} + Post_t + \mu_i + \epsilon_{it} \tag{3}$$

where i denotes a non-subsidized patent holder and t denotes the year. N_{it} is the number of new granted patents of the firm i in year t. $Trade_{it-1}$ is the number of patents that are sold from i in year $t-1$. $Post_t$ is a dummy, which equals to 1 if the year is after 2008. It captures the nation-wide trend of the new granted patents. μ_i is the firm fixed effect and ϵ_{it} is the error term. γ_0 is coefficient associated with $Trade_{it-1}$, and we expect a positive γ_0.

Both the $Trade_{it-1}$ and N_{it} can be affected by an unobserved firm-year level shock. We solve the endogeneity issue by instrumenting the $Trade_{it-1}$ using $IVTrade_{it-1}$. We define $IVTrade_{it}$ as

$$IVTrade_{it} = TotTrade_t \times X_{it-1} \tag{4}$$

where $X_{i,t-1}$ is the share of patents that are sold to firms with lower than 6 IPRs from i in year $t-1$, and $TotTrade_t$ is the total trade of patents within subsidized firms. The identification assumptions of this IV are as follows. First, the share $X_{i,t-1}$ is persistent and does not directly correlate with N_{it}. Since firms with lower than 6 IPRs have stronger incentives to purchase patents, i may sell more patents in year t, if $X_{i,t-1}$ is higher. Second, we assume that the trade within the subsidized firms, $TotTrade_t$, does not directly correlate with N_{it}.[21]

The first two columns of Table 5 show the results, where we report the OLS estimates and IV estimates respectively. We find that on average, the non-subsidized firms get around 2.7 more patents after 2008 (coefficient of $Post$). If they can sell 1 more patent, the increase of their new patents is around 1.312 to 3.837.

We are also interested in the value of new granted patents for non-subsidized firms. We measure the value by the value of new granted patents by renewal rates 3 years after granted and estimate a similar equation as (3). We use the similar IV as Column 2, assuming that $X_{i,t-1}$ and $TotTrade_t$ do not directly correlate with the renewal rate.

In the third and fourth column of Table 5, we report the OLS and IV estimates of the renewal rates. First, we find that after 2008, the renewal rates decline by around 9.6% on average (coefficient before $Post$). However, without an IV, we do not find a significant correlation between the change of trade and the change of renewal rates. In the last column (the IV estimate), we find that if a non-subsidized firm sells 1 more patents, the renewal rates of its new patents will decline about 0.6%. In other words, non-subsidized firms apply more low quality patents due to the subsidy.

To summarize our finding in this section, we show that due to the subsidy policy, more low value patents are granted, and more low quality patents are traded. In the next section, we build an equilibrium model, which is consistent with the empirical observations in this section, to quantify the aggregate welfare effect of the subsidy policy.

[21] In the first stage, we regress $Trade_{it} = \theta IVTrade_{it} + Post_t + \mu_i + error_{it}$, where we find that $\theta = 0.003$, and significant at the 1% level. The F value is around 101.44.

5 Model

We model the innovation market under the subsidy policy. People first innovate new patents and then patents are allowed to trade. The subsidy policy, which only counts the number of patents, can change the aggregate welfare either by distorting the decision of creating new patents or the trade of patents. Our model highlights three channels that the policy can affect the welfare. First, assuming the marginal cost of creating a patent is increasing in the patent's quality, people will shift to increase low quality patents under the subsidy. Second, the negative surplus trade will arise. Patents may be re-allocated to firms who cannot run them efficiently. Third, there is an echo effect from the trade to the innovation. Facing a high demand of low quality patents, people, who do not have chances to be subsidized, will invent more low quality patents to increase the chance to trade. In the following, we first lay our the model and then analyze how these three channels jointly determine the patterns in the data.

5.1 Environment

The economy has two types of agents: a mass I homogeneous patent inventors (which we call as inventors) and measure mI firms which can be subsidized if they have one patent (which we call as firms). The firms are heterogeneous in terms of productivity z. There are two differences between an inventor and a firm. First, the firm runs a patent with productivity z, while the inventor runs a patent with a homogeneous productivity which we normalize to be 1. We assume z follows a Bernoulli distribution: with probability G_H, $z = z_H$ and with probability $G_L = 1 - G_H$, $z = z_L$, and $z_H > 1 > z_L$. Second, the firm can get subsidized as long as it has a patent, and the inventor never has chances to be subsidized.

The economy has four stages. In the first stage, both of these two types of agents try to invent a patent. However, the firm faces a random shock: with probability σ, the firm has the chance to innovate by itself. With probability $1-\sigma$, the firm has no chance to innovate by itself and needs to purchase the patent from an inventor in the next stage.

We denote the patent return as x (we call it the patent quality), which is drawn from a Pareto distribution with the shape parameter $\mu > 1$ and the support $[\lambda, +\infty)$. Let θ denote the probability that the invention activity turns to be a patent successfully. Then the inventor or the self-innovation firm can choose λ and θ by paying an invention cost $C(\lambda, \theta)$. Thus with probability θ, the inventor or the self-innovation firm can have a patent, which is drawn from a Pareto distribution with the support $[\lambda, +\infty)$. We assume $C_\lambda > 0, C_\theta > 0$, $C_{\lambda\lambda} > 0$, $C_{\theta\theta} > 0$ and $C_{\lambda\theta} > 0$. The first two assumptions suggest that the innovation cost is increasing with quality and the succes rate. The next two inequalities suggest that the cost is convex. The last assumption suggests that the marginal cost of increasing θ is higher for high quality patent. We denote λ_i and θ_i as the minimum quality and success probability of the inventor, while denote $\lambda_f(z)$ and $\theta_f(z)$ as those of the firm z's choices. We denote the distribution of the patent quality distributions of these two types of agents as $H_i(x)$ and $H_f(x)$.

After the innovation stage, the trade of patents starts. We assume that sellers are inventors with patents (the measure of is θ_i), and buyers are firms which do not have chances to innovate in the first period (the measure is $m(1-\sigma)$). In other words, we do not allow inventors to purchase patents from other firms (i.e. z_H firm purchases patents from z_L firm). We also do not allow self-innovation firms, which fail to innovate patents successfully in the first stage (with a measure of $\sigma \int \theta_f(z)\, dG(z)$), to buy patents in the second stage. These assumptions could help us to get a much clean result.

In the third stage, the firm z gets subsidy $T(z)$ if it has a patent. We assume T is increasing in z to capture the idea that the benefit of the tax cut is greater if the firm is more productive (i.e. $T_H > T_L$). Notice that the subsidy does not depend on the quality of the patent, but only on whether the firm can get a patent or not. After the subsidy, all patent holders (including firms and inventors) can decide whether to use patents to produce or not. If a patent is used in the production, a renewal cost c must be paid in advance. Otherwise, the patent is expired.

Finally, in the fourth stage, the production starts. For those without active patents, the production=0. If the patent with quality x is owned by the firm z, the profit is $\pi_f(z, x) = \max(zx - c, 0) + T(z)$. The max operator suggests that if x is too low, the firm will not renew it and the production is 0. However, since the subsidy is awarded before the renew decision, the firm can always get $T(z)$. If the patent is owned by an inventor, the profit is $\pi_i(x) = \max(x - c, 0)$.

5.2 Renewal, Trading and Innovation Decisions

We first discuss the firm z's self-innovated decision, the problem is

$$\max_{\lambda_f(z),\theta_f(z)} \theta_f(z) \int \pi_f(z, x)\, dH_f(x) - C(\lambda_f(z), \theta_f(z)) \tag{5}$$

where the first term is the expected profit of the innovation and $C(\lambda_f(z), \theta_f(z))$ is the cost of innovation. The firm z chooses $\lambda_f(z)$ and $\theta_f(z)$ to maximize the net profit.

For notation easy, denote $\tilde{\lambda} = \lambda^{\mu}$. The cost function is denoted as $C\left(\tilde{\lambda}, \theta\right)$. We have the following result.

Proposition 1. If the Hessian matrix of the cost function $C\left(\tilde{\lambda}, \theta\right)$ is negative definite, and $C_{\tilde{\lambda}\theta}$ large enough, then when T increases, θ_f will increase, and λ_f will decrease.[22]

The proof of all the propositions is shown in Appendix C. The intuition is as follows:

[22] When we say T increases, it means that the subsidy increases for any z.

the benefit of a patent comes from the quality choice $\tilde{\lambda}_f$ and the chance to get a patent θ_f. However, the subsidy only depends on whether the firm has a patent or not. A large $C_{\tilde{\lambda}\theta}$ suggests that the marginal cost to improve the θ_f for high quality patents is very large. Hence a stronger subsidy gives a stronger incentive for the firm to increase θ_f, but decreases the quality.

We then switch to the patent trade market. We assume that the trade market is competitive. We denote the surplus of the trade between the inventor x and the firm z as

$$S(z,x) = \max(zx - c, 0) + T(z) - \max(x - c, 0) \tag{6}$$

So the value of the trade comes from two sources: the firm z can run the patent with different productivity; and the firm z can get subsidized if it has the patent. The trade will happen iff the surplus is positive.

If the patent is sold to $z_L < 1$, there is a profit loss ($z_L x - x < 0$). The larger x it is, the greater the loss is. Hence the z_L firm prefers to purchase low type patents to minimize the profit loss. On the other hand, if the patent is sold to $z_H > 1$, there is a profit gain ($z_H x - x < 0$). Hence the z_H firm will purchase high quality patents.

How does the subsidy decrease the welfare by distorting the patent trade market? When there is no subsidy ($T(z) = 0$), the z_L firm will not purchase any patent since there is no gain from the trade. Meanwhile, to get a positive surplus, patents sold to z_H firms must have quality greater than $\frac{c}{z_H}$. With the subsidy, z_L firms may purchase patents because the subsidy increases the surplus of the trade. Thus some inefficient rade arises. We can formally state the result as follows:

Proposition 2. Without the subsidy, there exists a cutoff $\hat{x}_H \geq \frac{c}{z_H}$. Patents whose $x \geq \hat{x}_H$ will be sold to z_H. All patents will be renewed if they are sold to z_H. No patents will be sold to z_L. With the subsidy, there exists two cutoffs $\hat{x}_H^{'}$ and $\hat{x}_L'$. Patents whose $x \geq \hat{x}_H'$ will be sold to z_H. And patents whose $x \leq \hat{x}_L'$ will be sold to z_L.

Before the subsidy, the renewal rate of the traded patents is 1. So after the subsidy, the renewal rate must decline. We have the following corollary.

Corollary 3. If T increases, the trade of patents will increase and the average quality of the traded patent will decline.

In the equilibrium, the patent trade market clears. Before the policy, we only need to consider z_H firm (z_L firm does not purchase any patent) and there are two possibilities. First, if the lowest traded patent generates 0 surplus ($S(z_H, \hat{x}_H) = 0$), the trade of patents below $\hat{x}_H$ cannot happen since it generates negative surplus. In this case, some z_H firms cannot buy patents since the total number of patents above $\hat{x}_H$ is lower than the number of buyers $m(1-\sigma)G_H$. In the second case, if $S(z_H, \hat{x}_H) > 0$, it means that the trade of patents below $\hat{x}_H$ can generate a positive surplus. Because all buyers have got patents above $\hat{x}_H$, patents below $\hat{x}_H$ will not be traded. Hence

$$1 - H_i(\hat{x}_H) \leq m(1-\sigma)G_H, \quad \text{with "} < \text{" if } S(z_H, \hat{x}_H) = 0 \tag{7}$$

where $1 - H_i(\hat{x}_H)$ is the number of patents that above $\hat{x}_H$.

After the subsidy, for the z_H firm, we have a similar equation as (7).

$$1 - H_i(\hat{x}_H') \leq m(1-\sigma)G_H, \quad \text{with "} < \text{" if } S(z_H, \hat{x}_H') = 0 \tag{8}$$

Meanwhile, patents below $\hat{x}_L'$ will be sold to z_L. We have two possibilities as well. First, if the surplus is 0 at the best patent sold to z_L ($S(z_L, \hat{x}_L') = 0$), it means that the trade of patents above $\hat{x}_L'$ will generate negative surplus. (Notice that $z_L < 1$, so that the greater the x is, the lower the trade surplus is). In this case, some z_L firms cannot buy patents. Meanwhile, if $S(z_L, \hat{x}'_L) > 0$, it means that all z_L firms can purchase patents and some positive surplus trade (patents higher than $\hat{x}'_L$) cannot happen. Thus we have

$$H_i(\hat{x}'_L) \leq m(1-\sigma) G_L, \quad \text{with " < " if } S(z_L, \hat{x}'_L) = 0 \tag{9}$$

where $H_i(\hat{x}'_L)$ is the number of patents below $\hat{x}'_L$.Equations (7) to (9) describe the equilibrium on the trade market.

Finally, we focus on the innovation of inventors. We assume that the surplus is separated between firm z and patent holder x with a share $1-\alpha$ and α. So let $z^*(x)$ denote the firm's productivity that will purchase x and let $W_i(x)$ denote the value of the inventor with a patent x after the trade. Then before the subsidy,

$$W_i(x) = \begin{cases} \alpha S(z^*(x), x) + \pi_i(x) & \text{if } x \geq \hat{x}_H \\ \pi_i(x) & \text{otherwise} \end{cases} \tag{10}$$

where the first line suggests that if the quality is above $\hat{x}_H$, the patent will be sold to $z^*(x)$ and generate a positive net surplus $S(z^*(x), x)$ and α share goes to the inventor. The second line suggests that if the quality is below than the cutoff $\hat{x}_H$, the inventor needs to run the patent by himself and gets profit $\pi_i(x)$. With the subsidy, $W_i(x)$ looks similar, but the condition in the first line in (10) changes to $x \geq \hat{x}'_H$ or $x \leq \hat{x}'_L$.

For the inventor, it chooses λ_i and θ_i to maximize the equation (10)

$$\max_{\lambda_i, \theta_i} \theta_i \int W_i(x)\, dH_i(x) - C(\lambda_i, \theta_i) \tag{11}$$

The equation suggests that the inventor tries to optimally choose the quality distribution and the success rate to maximize the expected payoff.

How does the subsidy change the inventor's innovation decision? Given the subsidy will increase the patent demand of z_L firms, the low quality patents will be more likely to sell. Hence inventors will decrease the quality of the invention and increase the possibility to get a patent. This result could be formally stated as the following proposition.

Proposition 4. If the Hessian matrix of the cost function $C\left(\tilde{\lambda}, \theta\right)$ is negative definite and $C_{\tilde{\lambda}\theta}$ is sufficiently large, then if T increases, θ_i will increase, λ_i will decrease.

We define an equilibrium as cutoffs $\hat{x}_H$, $\hat{x}'_H$ and $\hat{x}'_L$, choices of $\lambda_i, \theta_i, \lambda_f(z)$, $\theta_f(z)$ before and after the policy such that (i) given cutoffs, the firm's and inventor's optimization problems (5) and (11) are solved; (ii) the cutoffs solve the equilibrium conditions of the trade market (7) to (9).

6 Calibration

In this section, we discuss quantitative implications of the model. First, we normalize the cost of renewing the patent c to be 1, suggesting that the model's unit value is 900 RMB.[23] Second, we calibrate $m = 0.07$ to match the relative number of subsidized firms to inventors. We calibrate the measure of inventors I to match the aggregate subsidy 3.4 billion RMB in the data. Third, we assume that the surplus is split half-half between the patent buyer and seller ($\alpha = 0.5$). We assume the shape parameter of the Pareto distribution of patent quality is $\mu = 2$.

For the inventor, we assume the innovation cost is

$$C_i\left(\tilde{\lambda}_i, \theta_i\right) = v_0\left(1 - \tilde{\lambda}_i\right)^{-v_1} \theta_i^{v_2+1} \quad \text{if } 1 > \tilde{\lambda}_i > 0$$

[23]The renewal fee needs to pay in every year. The fee depends on the age of the patents. For patents whose age are lower than 3 years, the renewal fee is fixed at 900RMB since 2001.

where $v_0 > 0$ is a constant to measure the level of the cost of innovation. $v_1 > 0$ and $v_2 > 0$ control the curvature of the cost function. The above cost function suggests that $\tilde{\lambda}_i < 1$. If the quality becomes very bad ($\tilde{\lambda}_i$ is very small), then the cost is still positive. If $\tilde{\lambda}_i = 0$, we assume the patent quality $x = 0$ for sure and does not follow a Pareto distribution. And the cost to generate θ_i patents with 0 value becomes to $v_0 \theta_i^{v_2+1}$.

For the firm z_H and z_L, we assume their innovation costs are

$$C_H\left(\tilde{\lambda}_{fH}, \theta_{fH}\right) = v_{0H}\left(1 - \tilde{\lambda}_{fH}\right)^{-v_{1f}} \theta_{fH}^{v_2+1} \quad \text{if } 1 > \tilde{\lambda}_{fH} > 0$$

$$C_L\left(\tilde{\lambda}_{fL}, \theta_{fL}\right) = v_{0L}\left(1 - \tilde{\lambda}_{fL}\right)^{-v_{1f}} \theta_{fL}^{v_2+1} \quad \text{if } 1 > \tilde{\lambda}_{fL} > 0$$

where we allow $v_{1,f}$, the marginal cost of improving the quality of patents, to be different between the firm and the inventor, but the same between z_H and z_L. The firm z_H and z_L are different in terms of the level of the innovation cost v_{0H} and v_{0L}. These assumptions are motivated by the fact that the subsidy focuses on selecting out firms with good research abilities. Similarly, if firms choose $\tilde{\lambda} = 0$, the patent quality will be 0 for sure. Following Hayashi (1982), we calibrate $v_2 = 1$.

For simplicity, we assume $z_L = 0$. Thus z_L firm will always choose innovating the worst quality of patents ($\tilde{\lambda}_{fL} = 0$) since for the profit of any patents is 0.[24]

Now there are 10 parameters in our model: (i) the productivity values z_H and its measure G_H; (ii) the share of firms relying on self-innovation σ (iii) the cost functions of innovations v_0, v_{0H}, $v_{0,L}$, v_1, $v_{1,f}$; (iv) the subsidy levels T_H and T_L. The parameters are jointly estimated by minimizing the distance between the model moments and the following 10 targeted moments: (1) the number of self-developed patents per firm/inventor before and after the subsidy (4

[24] In the model, we can not identify the marginal cost of improving patents' quality of z_L firms (v_{1fL}), since they always chooses $\tilde{\lambda}_{fL} = 0$. This is another reason why we assume v_{1f} is the same within firms.

moments); (2) the average renewal rates 3 years after granted of firms and inventors before and after the subsidy (4 moments); (3) the number of traded patents per firm before and after the subsidy (2 moments).[25] Intuitively, (1) and (2) shed light on parameters of the innovation cost and the subsidy levels T_H and T_L. Meanwhile, the number of trade before the subsidy tells us σ, while the number of trade after the subsidy tells us G_H.

Table 6 lists the estimated parameters and reports the model fit by comparing the model moments and the moments we target. We normalize the annual number of granted patents per subsidized firm after the policy to be 1. All moments are well matched. As we can see, renewal rates of both firms and inventors decline after the policy is implemented, suggesting that the value of the patents declines. Second, the number of patents (traded or self-innovated) increases after 2008. All of these data facts suggest that the policy brings a negative welfare impact and we will quantify this effect in the next section.

In terms of the parameters, we find that $v_{1,f} = 0.132$ and $v_1 = 0.620$. So the subsidized firms have a lower marginal cost to improve the quality of patents. Meanwhile, $v_{0,H} = 0.023$ and $v_{0,L} = 0.190$. Hence z_H firms have a lower innovation cost than z_L firms. But both of their innovation costs are lower than inventors since $v_0 = 0.300$. Second, we find that about 8.2% firms are productive firms (z_H), and their productivity is 1% higher than inventors. Finally, we find that $T_H > T_L$, suggesting that more productive firms get more benefits from the subsidy program.[26]

[25] We use the 2000 to 2006 data sample to compute moments before the policy, and use 2008 to 2012 data sample to compute moments after the policy. The details of the calibration is provided in Appendix C.

[26] Notice that we do not restrict $T_H > T_L$ when we calibrate the model.

7 Results

7.1 The value of patents

The patents' value is described by the distribution of x, which is a Pareto distribution with shape parameter μ. However, the lower bound of the distribution is determined by the inventor. Panel A of Table 7 reports the average of the x before and after the subsidy policy. As we can see, the total average quality of patents declines from 1,341 RMB to 1,020 RMB. Thus the quality declines about 321 RMB for each patent on average.

7.2 Welfare decomposition

We now quantify the welfare effect of the subsidy policy. We allow the possibility that patents may generate knowledge spillovers. The aggregate welfare of the economy is defined as the sum of aggregate output of inventors and the firms, and the externality value generated from innovations. Formally, the aggregate welfare, denoted as U, is defined as

$$\begin{aligned} U &= \theta_i \int S\left(z^*\left(x\right), x\right) dH_i\left(x\right) + \theta_i \int \pi_i\left(x\right) dH_i\left(x\right) - C\left(\lambda_i, \theta_i\right) \\ &\quad + m\sigma \int [\int \theta_f\left(z\right) \pi_f\left(z, x\right) dH_f\left(x\right) - C\left(\lambda_f\left(z\right), \theta_f\left(z\right)\right)] dG\left(z\right) \\ &\quad - Subsidy + \varepsilon K \end{aligned} \tag{12}$$

where the first line is the total value of patents created by the inventors, which equals to the value of the traded patents plus the value of the non-traded patent minus the cost of innovation. The second line denotes the value of patents created by firms. In the third line, the total subsidy is defined as

$$Subsidy = \theta_i \int_{x \geq \hat{x}'_H} T_H dH_i\left(x\right) + \theta_i \int_{x \leq \hat{x}'_L} T_L dH_i\left(x\right) + m\sigma \int \theta_f(z) T\left(z\right) dG\left(z\right) \tag{13}$$

where the first two terms in (13) is the subsidy to firms purchasing patents, and the last term is the subsidy to self-innovation firms.

Finally, the last term K is the aggregate knowledge capital in the economy, which is defined as the total number of patents, adjusted by the quality.

$$K = \theta_i \int x dH_i(x) + m\sigma \int \int \theta_f(z) x dH_f(x) dG(z) \tag{14}$$

ε captures the spillover of the aggregate knowledge capital.[27] At the first place, we assume $\varepsilon = 0$, so that the patent does not have any externality. In the next subsection, we will allow positive ϵ.

In Panel B of Table 7, we show the aggregate welfare before and after the policy in billion RMB. The total subsidy increases from 0 to 3.4, and the aggregate output increases from 2.509 to 3.674. However, the aggregate profit (total output - innovation cost + subsidy) only increases by 2.869, from 1.255 to 4.124. The increase of the profit is less than the subsidy. This is because a huge amount of innovation cost is wasted on low quality patents, and the patent trade market efficiency goes down. Depending on ϵ, it may create the decline of the aggregate welfare. Assuming $\epsilon = 0$, the the aggregate welfare declines by 0.531 billion RMB, which is about 42% of the welfare before the subsidy. Furthermore, Since more patents are innovated, the aggregate knowledge capital increases from 0.621 to 1.413.

The welfare effect of the subsidy policy comes from three channels: (1) the innovation decisions of the subsidized firms θ_f and λ_f are different due to the subsidy; (2) the matching pattern $z^*(x)$ are different; (3) and there is a general equilibrium effect that since the (2) is different, it will further changes the θ_i and λ_i. We then decompose the welfare change

[27] There is a large literature estimate the knowledge spillovers of patents (See Bottazzi and Peri (2003) and Nelson (2009) for examples). The estimates suggest that if the aggregate number of patents increases by 1%, the aggregate output will increase by 0.2% to 0.4%.

as follows. First, we call the change of the second line after the policy is introduced as the direct effect on the innovation decision. Formally, it is defined as the change of

$$\Delta\left[m\sigma\int[\int\theta_f(z)\,\pi_f(z,x)\,dH_f(x)-C(\lambda_f(z),\theta_f(z))]dG(z)-m\sigma\int\theta_f(z)\,T(z)\,dG(z)\right]$$

where the first integral is the value of the self-innovated firms and the second integral is the subsidy received by those firms.

Second, we we define the welfare change from reallocation as the welfare change if the inventor's innovation decisions θ_i and λ_i do not change. In other words, the welfare change from the reallocation channel captures the welfare loss from inefficient trade, while assuming that the inventor does not change θ_i and λ_i after the subsidy. Define $\hat{x}_H^R$ and $\hat{x}_L^R$ as the cutoffs from the trade market equilibrium conditions (8) and (9).[28] Then the reallocation effect is defined as the change of

$$\Delta\left[\begin{array}{c}\theta_i\int S(z^*(x),x)\,dH_i(x)+\theta_i\int\pi_i(x)\,dH_i(x)-C(\lambda_i,\theta_i)\\ -\theta_i\int_{x>\hat{x}_H^R}T_H dH_i(x)-\theta_i\int_{x<\hat{x}_L^R}T_L dH_i(x)\end{array}\right] \tag{15}$$

where θ_i, λ_i and H_i denote the innovation decisions of the inventor before the subsidy policy. We do not change them in the equation (15). But $z^*(x)=z_H$ if $x>\hat{x}_H^R$ and $z^*(x)=z_L$ if $x<\hat{x}_L^R$. The second line in the equation (15) is the subsidy in this case.

Finally, the residual change of the aggregate welfare comes from the change of θ_i and λ_i, which we call it as the general equilibrium (GE) effect of the policy. The decomposition results are shown in Panel C of Table 7. All numbers are in billion RMB. Notice that our decomposition method will impose the sum of these three effects equal to to the total

[28] Notice that in the distribution H_i, we use the parameters λ_i and θ_i (parameters before the subsidy). However, when we evaluate the surplus $S(z,x)$, we include the subsidy T_H and T_L. Then the change of trade cutoffs is only because the subsidy, but does not depend on the change of λ_i and θ_i.

welfare decline (0.531). Our result suggests that the direct effect and GE effect contributes to the welfare decline by about 47% and 52%, while the contribution of the reallocation effect is tiny. In other words, the welfare effect of distorting inventors' innovation decisions is substantial, although they are directly subsidized in the InnoCom program.

Why the contribution of the reallocation channel is small? After the subsidy, some low value patents will be sold to z_L. If patents' x are lower than c, patents will never be used disregard whether they are sold or not. If patents' x are higher than c, they will be used in production if held by inventors, but not be used if owned by z_L firms. There are output loss only in the second case. However, since most patents that are sold to z_L are very low quality patents, the loss from the reallocation channel is small.

7.3 Optimal Subsidy Policy

A key problem of the subsidy policy is that the government does not observe the quality of patents, but only counts the number. In this subsection, we want to quantify the welfare change if the government can subsidize firms conditional on their patents' quality.

We focus on the following policy that the firm can get subsidize if its self-developed patent's quality is above $\bar{x}_s$ or if its purchased patent's quality is above $\bar{x}_t$.[29] The government now chooses $\bar{x}_s$ and $\bar{x}_t$ to maximize the social welfare, taking the individual firm and inventor's optimal decisions (equations (5) and (11)) and the trade market clearing conditions (equations (7) and (8)) as restrictions (hence the government solves a Ramsey problem).Meanwhile, we assume that the government has a deep pocket, so that the marginal

[29]For instance, the profit of the self-developed firm $z \in \{z_H, z_L\}$ is

$$\pi_f(z,x) = \begin{cases} \max(zx-c,0) + T(z) & \text{if } x > \bar{x}_s \\ \max(zx-c,0) & 0 \end{cases}$$

social value of the subsidy is 0.[30]

Proposition 5. If $m\,(1-\sigma)$ is small enough, (1) the optimal solution of the planner is to set $\bar{x}_s$ to be

$$\bar{x}_s = \left[\frac{\tilde{\lambda}_{fH}^{o\left(1-\frac{1}{\mu}\right)} \left(G_H \frac{T_H^2}{\theta_{fH}^{o(v_2-1)}} + G_L \frac{T_L^2}{\theta_{fL}^{o(v_2-1)}} \right)}{\varepsilon \frac{\mu}{\mu-1} \left(G_H \frac{T_H}{\theta_{fH}^{o(v_2-1)}} + G_L \frac{T_L}{\theta_{fL}^{o(v_2-1)}} \right)} \right]^{\frac{1}{\mu}} \tag{16}$$

where variables with subscript o denotes the values under the optimal subsidy.

(2) If $[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1} - \alpha T_H - (z_H-1)\tilde{\lambda}_i^o]\hat{x}_H^{o1-\mu} + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}} > 0$, then the planner sets $\bar{x}_t \leq \hat{x}_L^o$. Otherwise, the planner sets $\bar{x}_t \geq \hat{x}_H^o$.

The proofs are provided in Appendix D. First, let us consider the choice of $\bar{x}_s$. Intuitively, the externality that are not internalized by firms' innovation is the knowledge spillover. So if $\varepsilon = 0$, we can see that $\bar{x}_s = +\infty$, implying that no self-innovated firms will be subsidized.

However, the inventor's problem is more complicated. There are three sources of externality. First, there is a classical hold-up problem making inventor's innovation lower than the social optimum. After making an innovation and selling it to a firm, the inventor can only get $\alpha < 1$ share of the surplus. Inventors have to make an innovation investment before finding the patent buyers, and a large innovation investment translates into higher gain of the trade. But the inventor cannot fully enjoy the gain, so they invest in innovation insufficiently. Second, if an inventor increases its innovation, it decreases other inventors' probabilities to trade with firms (i.e pushing up the cutoff $\hat{x_H}$). However, the inventor does not internalize this negative externality and its innovation may be too much comparing to the social optimal case. Lastly, the knowledge spillover externality exists when $\epsilon > 0$. So that inventor may under-invest on the innovation. If the first and third effects dominate the second effect, the inventor's innovation is too low comparing to the optimum. The planner will choose a low

[30]The setup of the problem is in Appendix D.

$\bar{x}_t$ so that more trade an be subsidized, hence give inventors stronger incentives to innovate. However, if the second effect is strong enough, the inventor's innovation is too much. The government does not want to encourage its invention, hence will choose a high (for instance, $\bar{x}_t = \infty$, then none of the trade will be subsidized).

The inequality condition in (2) of the previous proposition evaluates which effect dominates. If α is greater (the first externality becomes weaker) or ε is smaller (the third externality becomes weaker), it is more likely that $[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1} - \alpha T_H - (z_H-1)\tilde{\lambda}_i^o]\hat{x}_H^{o1-\mu} + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}} < 0$. In this case, the second externality is more likely to dominate the first and the third externality. Hence the government tends to decrease the subsidy by setting a high minimal quality requirement $\bar{x}_t$.

Column 1 of Table 8 shows the welfare effect under the optimal policy, in which we assume $\varepsilon = 0$. First, not surprisingly, the government does not subsidize any self-developed patents ($\bar{x}_s = \infty$) since the externality of knowledge capital does not exist. Second, the government does not want to subsidize any patents from trade as well ($\bar{x}_t = \infty$). This means that under our parameters, the second externality is strong enough so the government does not want to encourage any invention from inventors. So we can see that the aggregate subsidy, knowledge capital, the net profit (total profit-subsidy) and the welfare are the same as those without any subsidy (the first column of Panel B in Table 7).

In Column 3, we consider the externality of the knowledge spillover. Following Jaffe (1986) and Bottazzi and Peri (2003), we calibrate $\epsilon = 0.1$, so that a 1% increase of the number of patents contributes the aggregate welfare the same as 0.2% increase of the aggregate output.[31] We then re-solve the optimal subsidy policies. In this case, the government wants to set $\bar{x}_s = 1{,}241$ RMB and $\bar{x}_t = 715$ RMB. The total subsidy is 0.215 billion RMB, much

[31] Before the subsidy, a 1% increase of the knowledge capital is about 0.05 (4.996*0.01), and results 0.05ϵ increase of U. A 0.2% increase of the aggregate output will increase U by about 0.005 (2.509*0.002). Hence $\epsilon = 0.1$ in our model.

lower than the subsidy which only counts the patent numbers (3.4). Comparing to the Column (1) and (3), the knowledge capital increases to 5.492 billion RMB, and the net profit slightly goes down. In the end, the aggregate welfare is 1.773 billion RMB, while the aggregate welfare without subsidy is 1.754 billion RMB (1.255+0.1*4.996).

Starting from 2016, Chinese government imposes restrictions on the subsidy of traded patents. There are two ways firms can acquire patents externally: purchasing patents directly or purchasing exclusive licenses. The second way usually is cheaper and faster. Before 2016, the government does not distinguish these two ways. In other words, patents that are owned under exclusive licenses are counted in the same way as directly purchased patents or self-innovated patents. Hence firms can easily acquire lots of low quality patents by signing exclusive licenses with others to be subsidized. To avoid it, exclusive-license-owned patents are not counted in the evaluation of the InnoCom program since 2016.

To evaluate the welfare effect of this policy change, we consider $\bar{x}_s = 0$ and the government only optimally chooses $\bar{x}_t$. Thus the government counts the number of self-developed patents without taking into account their quality, but only subsidize high quality external purchased patents. Since $\bar{x}_t$ is optimally chosen, the welfare effect in our exercise could be considered as the up-bound of the 2016 policy. The results are shown in Column 2 of Table 8. Again we assume $\varepsilon = 0$. In this case, the government does not want to subsidize any purchased patents ($\bar{x}_t = \infty$) because the second externality effect is strong enough. The government only subsidizes self-developed patents and the total subsidy is about 0.264 billion. The aggregate knowledge capital is about 4.337 billion. However, since the subsidized firms create too much low quality patents, the net profit and the aggregate welfare (1.004) are lower than those without subsidy. In Column 4, we re-do the similar exercise, but assume $\varepsilon = 0.1$. As we can see, the government only counts the number of purchased patents if the quality is higher than 715 RMB. The aggregate welfare is 1.441 billion RMB, lower than that without subsidy (1.754).

How does the welfare of optimal policy depend on ε? Figure 6 shows the results under different ϵ. We range ϵ from 0 to 0.5, and solve the optimal cutoffs $\bar{x}_s$ and $\bar{x}_t$.[32] The first graph shows the aggregate subsidy under the optimal policy (solid line) and the subsidy when only counting the number of patents (dashed line). As we can see, the optimal subsidy level is increasing from 0 to 1.05 billion RMB, when ϵ increases. But it is always much lower than the subsidy when only counting the number of patents (3.4 billion RMB). The second graph shows the aggregate knowledge capital under the optimal policy (solid line), and the knowledge capital without subsidy or under the counting-number subsidy (two dashed lines). We can see that the optimal knowledge capital increases in ϵ. The third graph shows the net profit (profit excluding the surplus). When ϵ increases, we can see the net profit under the optimal policy declines. This is because that the externality of K becomes more important, so the government tries to encourage more innovations. So the private marginal net profit of the innovation is lower than the marginal innovation cost. The fourth graph shows the aggregate welfare in the optimal policy (solid line), and the welfare without the subsidy or under the counting-number subsidy (two dashed lines). The two dashed lines will increase when ϵ increases due to the value from the knowledge capital increases. We can see that when $\epsilon \geq 0.4$, the current subsidy scheme (only counting the patent number) can generate a greater welfare than the economy without any subsidy. However, the optimal policy always dominate the other two cases. The last graph shows the choice of $\bar{x}_s$ (solid line) and $\bar{x}_t$ (dashed line). From Table 8, we know that both of them are infinity when $\epsilon = 0$. So the graph starts from $\epsilon = 0.1$. When ϵ increases from 0.1 to 0.5, $\bar{x}_s$ declines from 1,241 RMB to about 800 RMB, because the government wants to increase the probability to get subsidy to encourage more innovation. At the same time, $\bar{x}_t$ is around 715 RMB.

[32]When $\epsilon = 0.5$, it means that the welfare change of 1% increase of K is equivalent to 1% increase of the aggregate output.

8 Conclusion

In this paper, we study how the patent subsidy policy, the InnoCom program, change the aggregate welfare. We show that the subsidy policy decreases the value of patents either through declining the quality of new granted patents or through the inefficient trade of patents. We decompose the welfare change into three channels: the direct innovation effect, the reallocation effect and the indirect effect from the trade market. We find that the most important driving force is the last channel, suggesting that it is crucial to jointly consider the patent trade market and the entry of new patent at the same time.

References

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.

Bartik, T. J. et al. (1991). Who benefits from state and local economic development policies? *Books from Upjohn Press*.

Barwick, P. J., M. Kalouptsidi, and N. B. Zahur (2019). China's industrial policy: An empirical evaluation. Technical report, National Bureau of Economic Research.

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. V. Reenen (2002). Do rd tax credits work? evidence from a panel of countries 1979–1997. *Journal of Public Economics 85*(1), 1 – 31.

Bottazzi, L. and G. Peri (2003). Innovation and spillovers in regions: Evidence from european patent data. *European economic review 47*(4), 687–710.

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

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

Das, S., M. J. Roberts, and J. R. Tybout (2007). Market entry costs, producer heterogeneity, and export dynamics. *Econometrica 75*(3), 837–873.

Deng, Y. (2007). The effects of patent regime changes: A case study of the european patent office. *International Journal of Industrial Organization 25*(1), 121–138.

Galasso, A., M. Schankerman, and C. J. Serrano (2013). Trading and enforcing patent rights. *The RAND Journal of Economics 44*(2), 275–312.

Gans, J. S. and S. Stern (2000). Incumbency and r&d incentives: Licensing the gale of creative destruction. *Journal of Economics & Management Strategy 9*(4), 485–511.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018). Bartik instruments: What, when, why, and how. Technical report, National Bureau of Economic Research.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica: Journal of the Econometric Society*, 213–224.

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.

Jaffe, A. B. (1986). Technological opportunity and spillovers of r & d: Evidence from firms' patents, profits, and market value. *The American Economic Review*, 984–1001.

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.

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

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

Liu, E. (2019). Industrial policies in production networks. *The Quarterly Journal of Economics 134*(4), 1883–1948.

Nelson, A. J. (2009). Measuring knowledge spillovers: What patents, licenses and publications reveal about innovation diffusion. *Research policy 38*(6), 994–1005.

Neumark, D. and H. Simpson (2015). Place-based policies. In *Handbook of regional and urban economics*, Volume 5, pp. 1197–1287. Elsevier.

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

Schankerman, M. (1998). How valuable is patent protection? estimates by technology field. *the RAND Journal of Economics*, 77–107.

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

Serrano, C. J. (2018). Estimating the gains from trade in the market for patent rights. *International Economic Review 59*(4), 1877–1904.

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

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.

Wilson, D. J. (2009). Beggar thy neighbor? the in-state, out-of-state, and aggregate effects of r&d tax credits. *The Review of Economics and Statistics 91*(2), 431–436.

Yi, F., C.-Y. L. Lawell, and K. E. Thome (2019). A dynamic model of subsidies: Theory and application to ethanol industry.

Zhao, C., X. Qu, and S. Luo (2019). Impact of the innocom program on corporate innovation performance in china: Evidence from shanghai. *Technological Forecasting and Social Change 146*, 103–118.

Tables and Figures

Table 1: The InnoCom Program in Beijing

Year	No. of New InnoCom firms	Number of Other Patent Holders	Subsidy per InnoCom firm (1 mil RMB)
2008	2,608	71,187	-
2009	5,167	76,432	0.578
2010	6,401	83,132	0.533
2011	7,204	90,670	0.744
2012	7,920	99,032	-
Average	5,860	84,091	0.618

Notes: This table shows the number of new InnoCom firms in Beijing from 2008-2012, and the number of patent holders, which are defined as firms, individuals, or institutions that ever granted a patent in Beijing. The last column shows the annual subsidy (tax return) per InnoCom firm, which is computed using the State Administration Tax Data. There are no InnoCom firms in the tax data in 2008 and 2012.

Table 2: Summary Statistics of the Subsidized Firms Covered by the InnoCom Program in 2008

	Obs.	Mean	Sd
A: The complete list of firms under InnoCom Program			
Sales (1mil RMB)	2,608	14.41	116.57
Assets (1mil RMB)	2,608	18.88	98.79
No. of workers	2,608	161	369
Research staff share	2,608	0.44	0.23
No. of IPRs	2,608	8.82	13.29
No. of patents	2,608	3.57	11.66
No. of software	2,608	5.16	7.49
Score	2,608	81.78	7.19
B: Firms under InnoCom Program matched with IPRs information			
Sales (1mil RMB)	1,776	19.04	140.65
Assets (1mil RMB)	1,776	25.21	118.82
No. of workers	1,776	188	407
Research staff share	1,776	0.40	0.22
Score	1,776	82.11	7.17

Notes: Panel A of the table shows the summary statistics of the subsidized firms covered by the InnoCom Program in 2008. Panel B shows the summary statistics of those subsidized firms matched with the IPRs information. Data source: Beijing Municipal Science and Technology Commission and Chinese State IPRs Office.

Table 3: The Number of New Patents in Relation to the Timing of InnoCom Policy

	(1)	(2)	(3)	(4)
T	0.655***	0.244		
	(0.059)	(0.244)		
$\tilde{T}$				1.308***
				(0.264)
$(IPR_{2007} < 6) \times T$		0.474**	0.269**	
		(0.248)	(0.133)	
$(IPR_{2007} < 6) \times \tilde{T}$				-1.487***
				(0.266)
$Patent_{it-1}$	0.109***	0.111***	0.077***	0.110***
	(0.010)	(0.010)	(0.003)	(0.010)
Firm FE	Y	Y	Y	Y
Year FE			Y	
Obs.	15,832	15,832	15,832	15,832
Adj. R-squared	0.294	0.295	0.326	0.290

Notes: This table presents the estimation results of the equation (1) using the patents information of the 2008 subsidized firms from 2001-2010. The observation is at firm-year level. The dependent variable N_{it} is the number of patents obtained by the firm i in year t. For Columns 1 to 3, $T = 1$ if year=2008 and 0 otherwise. For Column 4, $\tilde{T} = 1$ if year=2007 and 0 otherwise. Standard errors are reported in the parentheses and clustered at the firm-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: The Results of the Renewal Rate Estimation

	(1)	(2)	(3)	(4)	(5)
$(T_{ij} = 2008)$	-0.009**	0.008			0.004
	(0.004)	(0.010)			(0.021)
$(T_{ij} = 2007)$				0.008	
				(0.017)	
$(T_{ij} = 2008) \times (IPR_{2007} < 6)$		-0.018*	-0.018*		-0.018
		(0.011)	(0.011)		(0.021)
$(T_{ij} = 2007) \times (IPR_{2007} < 6)$				-0.011	
				(0.017)	
$(T_{ij} = 2008) \times IPR_{2007} \times (IPR_{2007} < 6)$					0.012**
					(0.005)
$(T_{ij} = 2008) \times IPR_{2007} \times (IPR_{2007} \geq 6)$					0.000
					(0.000)
Z_{ij}	-0.005	-0.005	0.004	-0.007	-0.004
	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)
Patent Age	-0.013***	-0.013***	-0.018***	-0.012***	-0.013***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
IPC-year FE	Y	Y	Y	Y	Y
Firm-year FE	Y	Y	Y	Y	Y
Transaction-year FE			Y		
Obs.	20,167	20,167	20,167	20,167	20,167
Adj. R-squared	0.209	0.209	0.211	0.209	0.201

Notes: This table presents the estimation result of the equations (2) using the patent renewal information of 2008 subsidized firms from 2001 to 2010. The observation is at firm-patent-year level. The dependent variable $V_{ijt} = 1$ if the patent j is renewed in year t and 0 otherwise. For the column 1 to 3, $(T_{ij} = 2008)$ equal to 1 if the firm i gets the patent j in year 2008 and 0 otherwise. For the column 4, $(T_{ij} = 2007)$ is 1 if the firm i gets the patent j in year 2007 and 0 otherwise. $Z_{ij} = 1$ if the firm i gets the patent j by acquisition and 0 if self invention. IPR_{2007} is the IPR number in 2007 of firm i. IPC is the 5-digit International Patent Classification code. Patent cohort FE is a sequence of dummies for T_{ij}. Standard errors are reported in the parentheses and clustered at the patent-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: The Results of Granted Patents of Inventors

	(1)	(2)	(3)	(4)
	No. of granted patents		Renewal rate 3 years after granted	
Trade	1.312***	3.837***	-0.002	-0.006*
	(0.275)	(0.761)	(0.002)	(0.003)
Post	2.725***	2.660***	-0.096***	-0.096***
	(0.205)	(0.214)	(0.008)	(0.008)
Inventor FE	Y	Y	Y	Y
IV		Y		Y
Obs.	51,099	51,099	51,099	51,099
Adj. R-squared	0.393	0.392	0.203	0.203

Notes: This table presents the estimation results of number and citation of new patents (equation (3)) using the non-subsidized firms from 2001-2015. The dependent variable of the first two columns is the number of new patents of a firm in a year. The dependent variable of the third and fourth columns is the average renew rates of new patents 3 years after granted. *Trade* is the number of patents selling from non-subsidized firms to subsidized firms. We instrument it using the initial trade to firms with less than 6 patents and the patents trade within subsidized firms in the second and fourth column. $Post = 1$ if the year is after 2008 and 0 otherwise.Standard errors are reported in the parentheses and clustered at the firm-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Parameters and Model Fit

Param.	Value	Moments	Data	Model
z_H	1.010	Renewal rate of the subsidized firms before subsidy	0.793	0.793
$v_{1,f}$	0.132	Renewal rate of the subsidized firms after subsidy	0.756	0.756
$v_{0,H}$	0.023	No. of granted patents per subsidized firm before subsidy	0.823	0.823
$v_{0,L}$	0.190	No. of granted patents per subsidized firm after subsidy	1.000	1.000
v_1	0.620	Renewal rate of inventors before subsidy	0.632	0.632
v_0	0.300	Renewal rate of inventors after subsidy	0.590	0.590
T_H	0.596	No. of granted patents per inventor before subsidy	0.384	0.384
T_L	0.410	No. of granted patents per inventor after subsidy	0.608	0.608
G_H	0.082	No. of trade per subsidized firm before subsidy	0.010	0.010
σ	0.878	No. of trade per subsidized firm after subsidy	0.122	0.122

Notes: This table reports the parameters and model fit by comparing the moments in the model and the data. The moments before and after the subsidy are computed using the data from 2000 to 2006, and 2008 to 2012 respectively. All moments are computed based on the balanced sample (i.e. the sample of subsidized firms and inventors are fixed the same before and after the policy).

Table 7: Impacts of the Subsidy Policy

A: Average quality per patent (in RMB)			
	Before Subsidy	After Subsidy	Change (after-before)
Ave. quality	1,341	1,020	-321
B: Aggregate welfare and subsidy (in bil. RMB)			
	Before Subsidy	After Subsidy	Change (after-before)
Aggregate output	2.509	3.674	1.166
Aggregate Profit	1.255	4.124	2.869
Subsidy	0.000	3.400	3.400
Knowledge capital	4.996	6.350	1.354
Welfare ($\epsilon = 0$)	1.255	0.724	-0.531
C: Welfare change decomposition (in bil. RMB)			
Welfare change ($\epsilon = 0$)	Direct effect	Reallocation	GE effect
-0.531	-0.251	-0.004	-0.276

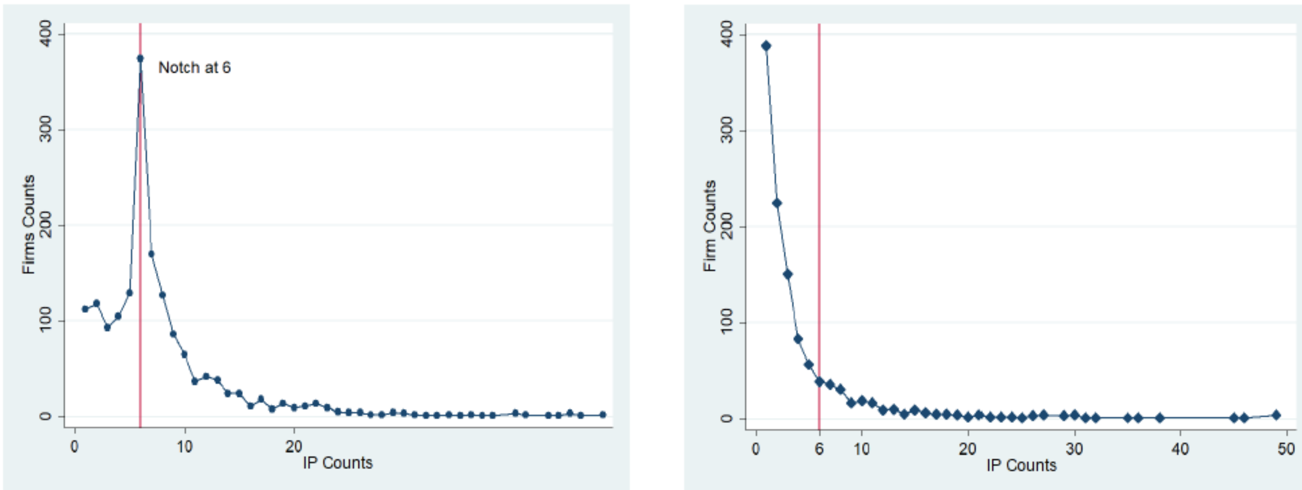
Notes: Panel A of the table reports the average of patent quality (x) before and after the subsidy. Panel B of the table reports the aggregate output, profit, subsidy, knowledge capital and welfare ($\epsilon = 0$) before and after the subsidy. The aggregate profit = output-innovation cost+subsidy. Panel C decomposes the welfare change into three channels.

Table 8: The Optimal Subsidy Policy

	$\epsilon = 0$		$\epsilon = 0.1$	
	(1)	(2)	(3)	(4)
$\bar{x}_s$ (in 1,000 RMB)	∞	0	1.241	0
$\bar{x}_t$ (in 1,000 RMB)	∞	∞	0.715	0.715
Subsidy (in bil. RMB)	0	0.264	0.215	0.296
Knowledge capital (in bil. RMB)	4.996	4.337	5.492	4.368
Net Profit (in bil. RMB)	1.255	1.004	1.224	1.004
Welfare (in bil. RMB)	1.255	1.004	1.773	1.441

Notes: This table shows results if the government requires minimum qualities for firms to get subsidized. Column (1) assumes $\epsilon = 0$ and the government optimally imposes two minimum qualities for self-developed patents and purchased patents. Column (2) assumes $\epsilon = 0$ and the government only optimally imposes a minimum quality for purchased patents. Column (3) assumes $\epsilon = 0.1$ and the government optimally imposes two minimum qualities for self-developed patents and purchased patents. Column (4) assumes $\epsilon = 0.1$ and the government only optimally imposes a minimum quality for purchased patents. Net profit=aggregate profit-subsidy=welfare+ϵ knowledge capital.

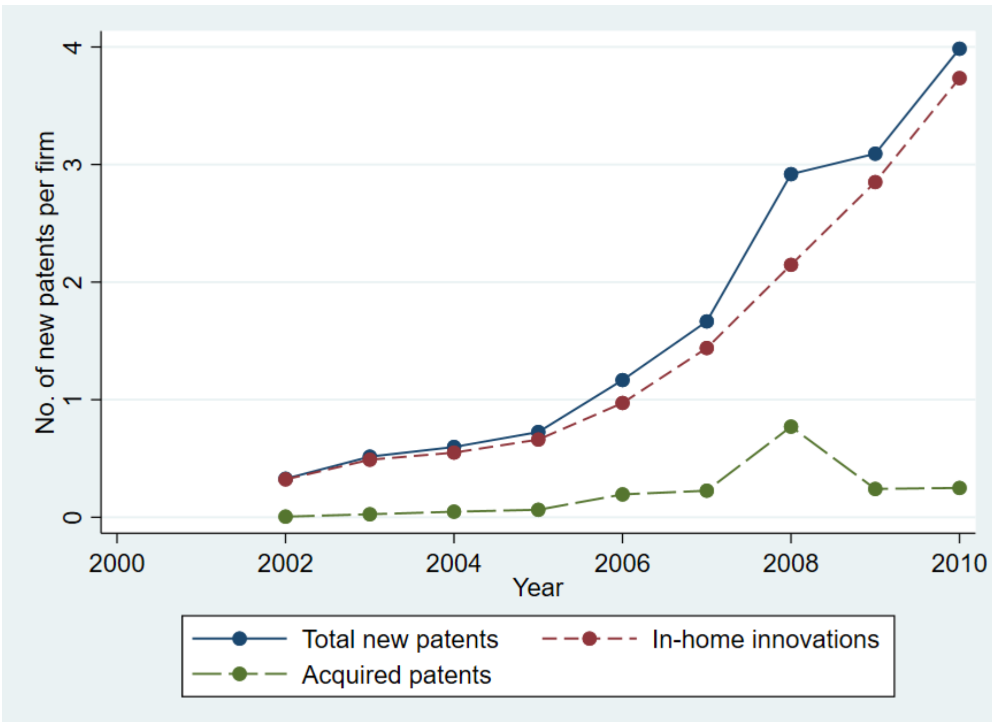
Figure 1: The IPRs Holding Distribution



Year 2008 Year 2007

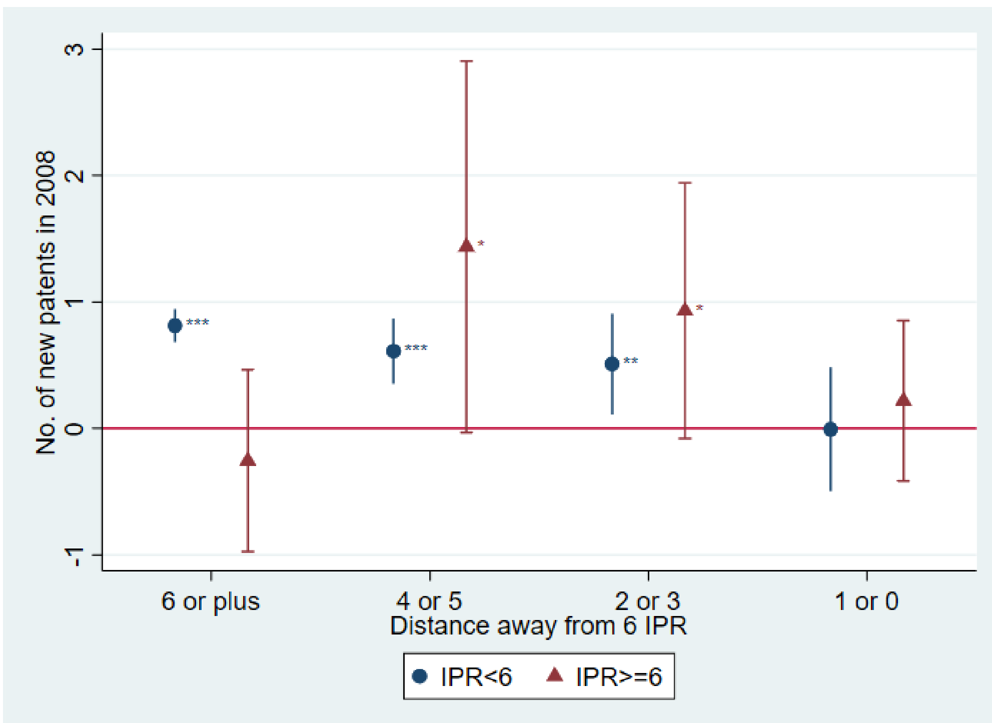
Notes: This figure shows the IPRs (patents and software copyrights) holding distributions of the subsidized firms in Beijing in year 2008 and 2007.

Figure 2: The Number of New Patents per Subsidized Firms Covered in 2008 InnoCom Program



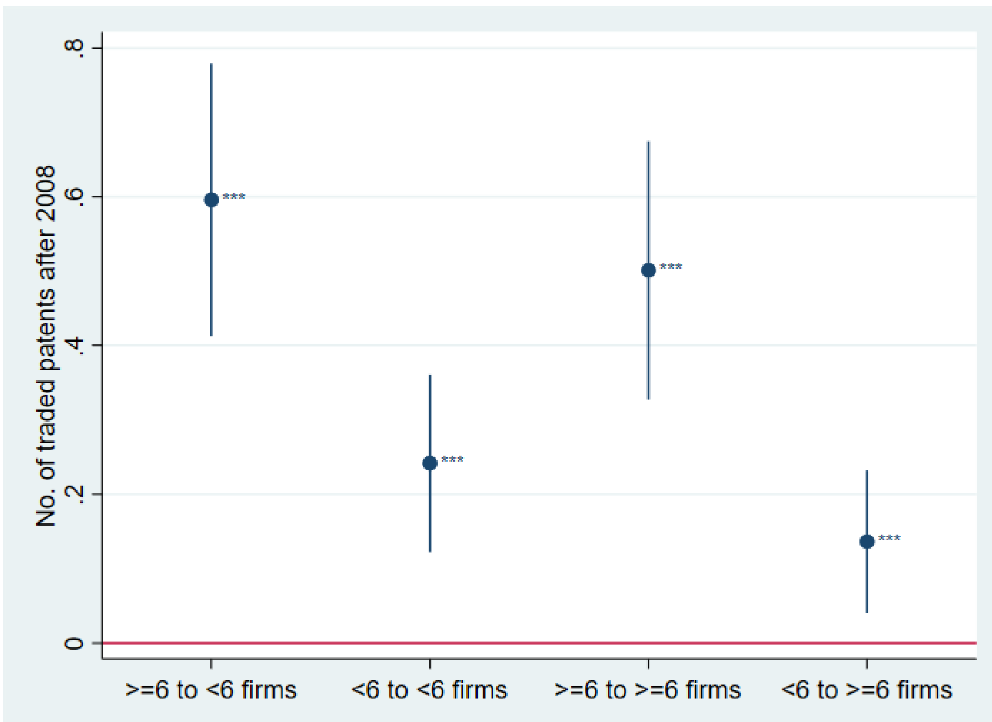
Notes: This figure shows the total number of new patents (solid), the number of self-developed patents (dash) and the number of purchased patents (long-dash) per subsidized firms covered by the 2008 InnoCom program from 2002-2010.

Figure 3: The Change of New Patents per Subsidized Firms in 2008



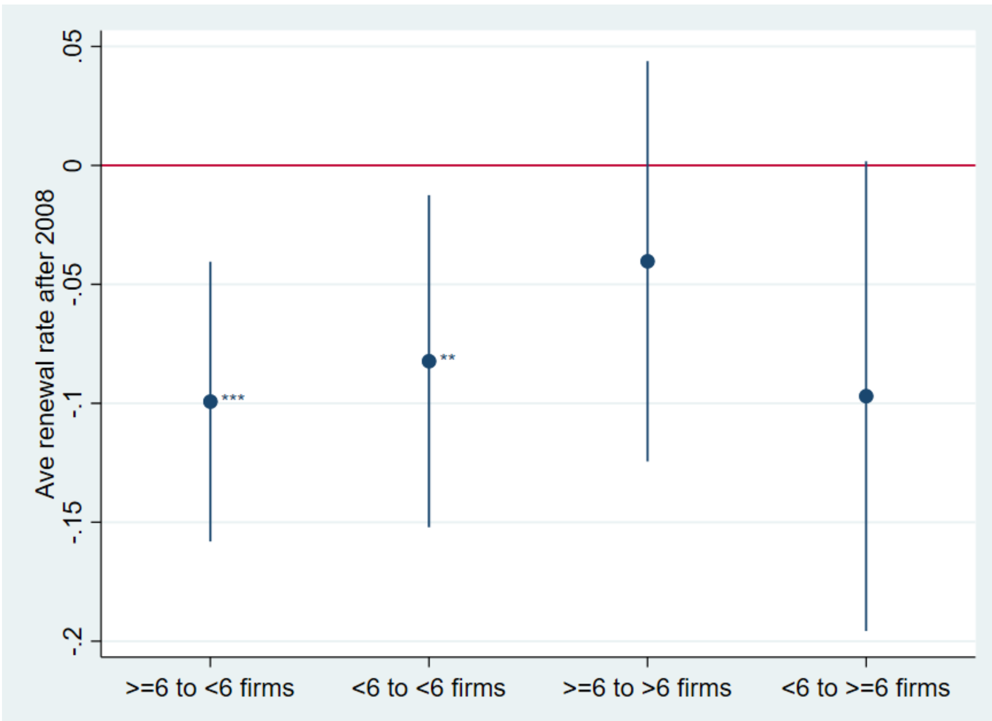
Notes: This figure shows the change of the new patents counts per subsidized firms covered by the 2008 InnoCom program in the year 2008. The x-axis denotes the distance of the number of IPRs away from 6 in year 2007. Standard errors are clustered at the firm-year level. 95% intervals are shown. *** p<0.01, ** p<0.05, * p<0.1.

Figure 4: The Number of Traded Patents after 2008



Notes: This figure shows the change of the traded patents counts after the year 2008. The x-axis denotes four groups of trade: firms with more than 6 IPRs sell to firms with lower than 6 IPRs; firms with fewer than 6 IPRs sell to firms with fewer than 6 IPRs; firms with more than 6 IPRs sell to firms with more than 6 IPRs; and firms with fewer than 6 IPRs sell to firms with more than 6 IPRs. Standard errors are clustered at the firm-year level. 95% intervals are shown. *** p<0.01, ** p<0.05, * p<0.1.

Figure 5: The Renewal Rates of Traded Patents after 2008



Notes: This figure shows the change of the renewal rate of traded patents 3 years after trade, after the year 2008. The x-axis denotes four groups of trade: firms with more than 6 IPRs sell to firms with lower than 6 IPRs; firms with fewer than 6 IPRs sell to firms with fewer than 6 IPRs; firms with more than 6 IPRs sell to firms with more than 6 IPRs; and firms with fewer than 6 IPRs sell to firms with more than 6 IPRs. Standard errors are clustered at the firm-year level. 95% intervals are shown. *** p<0.01, ** p<0.05, * p<0.1.

Figure 6: The Impacts of Optimal Subsidy Policy Under Different ϵ

subsidy-eps-converted-to.pdf

KK-eps-converted-to.pdf

netprofit-eps-converted-to.pdf

welfare-eps-converted-to.pdf

xbar-eps-converted-to.pdf

Online Appendix (not for publication in print)

A Details of the InnoCom Program

This section has two parts. In the first part, we explains the details on how the InnoCom program evaluates firms. In the second part, we compares the program with other innovation subsidy programs in China.

Table A1 lists the conditions that the firm must satisfy in order to apply for the Innocom program. As we can see, the pre-requirements are based on the R&D-to-sales ratio, the sales of high-tech products, the share of R&D workers, and the number of IP rights (IPRs).

After the application, the local government evaluates the firms based on four parts:the intellectual property rights, the ability to manage R&D, S&T commercialization, and the firm growth. The full marks in the 4 parts are 30, 30, 20, 20 respectively. We explain the grading criteria of the last three categories in Table A2. We can see that for each category, it has 6 levels (A to F). Different levels represent different points the firm can get in that category. The total points of the three categories are 70.

The other 30 points come from the holding of the intellectual property rights. We explain the grading criteria in the first column of Table A3. A noteworthy point is the grading scheme of the IPRs. The firm gets an "A" if they have 6 or more IPRs. Here we emphasize two important points. First, the subsidy program only counts the number of IPRs but does not take into account the quality of the IPRs. Second, the firms do not get any extra points if they have more than 6 IPRs.

Moreover, when the program was first introduced in 2008, the program recognized two ways of holding an IPR: the firm can either be the owner of the IPR (either innovate in-house or purchase from outside) or can hold the IPR under an exclusive license. In other words, for those patents which the firm borrow from others, the program still count them.

In 2016, there was a reform in the grading criteria. Other three parts do not change, but the grading criteria of the IPRs holding changed dramatically. In the second column of Table A3, we explain the new criteria. Comparing these two criteria, there are a few points to highlight. First, the IPRs borrowed from outside is counted since 2008 but not counted after 2016.

Table A1: Necessary Requirements of the Innocom Program

R&D Intensity	6% if sales < 50 millions of RMB 4% if sales > 50 & sales < 200 millions of RMB 3% if sales > 200 millions of RMB
Sales of High Tech Products	60% of total sales
Workers with College Degree	30% of workforce
R&D Workers	10% of workforce
IP right	Necessary

B The Data Matching Procedure

We use firm names to match different data sets. To clean and standardize firm names, we follow these steps: 1. We trim all special symbols and punctuation marks that are not letters, characters or numbers. 2. We remove various corporate form, such as "limited corporation" or "subsidiary". 3. We convert all full-width letters and numbers into half-width ones. 4. All lower cases are changed in to upper cases.

After getting the standard firm names, we then merge the patent assignment data with the name list of subsidized firms. To see the performance of our matching process, Table A4 shows the matched rate of subsidized firms covered by the InnoCom program in 2008 in Beijing and firms in the IPR (patent and software) assignment data. We separate the firms into two groups: firms with patents and firms without patents (only have other IPRs). First, the matched rate between subsidized firms and patent assignment data is as high as 94%. This ratio is comparable to Hu and Jefferson (2009). However, the matched rate is only 36% between subsidized firms and software assignment data. This may be because that those subsidized firms without patents have other non-software IPRs.

Table A2: The Grading Scheme of the Innocom Program

Category	Grading Criteria
Ability to transform research projects to business (30 points)	A. Above 4 items (30-24 points) B. 3-4 items (30-24 points) C. 2-3 items (18-12 points) D. 1-2 items (12-6 points) E. 1 item (0-6 points) F. 0 item (0 point)
Research management ability (20 points)	(1) Make research proposals (2) Have R&D expenditure accounting system (3) Co-operate with research institutions (4) Have independent research departments (5) R&D workers' performance evaluation system A. 5 items (16-20 points) B. 4 items (12-16 points) C. 3 items (8-12 points) D. 2 items (4-8 points) E. 1 item (0-4 points) F. 0 item (0 point)
Firm growth (20 points)	Asset growth (10 points) A. $\geq$35% (8-10 points) B. $\geq$25% (6-8 points) C. $\geq$15% (4-6 points) D. $\geq$5% (2-4 points) E. $<$ 5% (0-2 points) F. ≤ 0 (0 points) Sales growth (10 points) A. $\geq$35% (8-10 points) B. $\geq$25% (6-8 points C. $\geq$15% (4-6 points) D. $\geq$5% (2-4 points) E. $<$5% (0-2 points) F. ≤ 0 (0 points)

Table A3: The Grading Criteria of the Innocom Program: IPRs

	2008	2016
Exclusively Licensed IP	permitted	Not permitted
Quantity Requirement	A.(30-24) 6 or 1 invention patent B.(24-18) 5 C.(18-12) 4 D.(12-6) 3 E.(6-0) 1-2	IP are divided into 2 types: invention patent, plant variety, crop variety, new drugs, first class Chinese medicine variety, integrated circuit layout design are classified as type I, utility model patent, design patent, Software Copyright as type II A.(7-8) 1 type I or more B.(5-6) 5 type II C.(3-4) 3-4 type II D.(1-2) 1-2 type II
Other difference		(0-8) The extent IP supports main business (0-8) How firm obtains an IP (self innovation scores more than assignment) (0-6) How advance the technology (0-2) Whether participate in creating industry standard

Table A4: Matched Rate of Subsidized Firms

	No. of subsidized firms	No. of matched firms	Matched rate
Firms with patents in 2007	1,448	1,355	94%
Firm without patents in 2007	1,160	421	36%

Notes: This table reports the matched rate of subsidized firms covered by the InnoCom program in 2008 in Beijing and firms in the patent assignment data.

C Solutions of the Model

This section presents the solutions of the model in detail.

C.1 Proof of proposition 1

The optimization of the problem (5) can be written as

$$\max_{\lambda_f,\theta_f} \mu\tilde{\lambda}_f\theta_f \int_{\max\left(\lambda_f,\frac{c}{z}\right)}^{+\infty} (zx-c)x^{-(\mu+1)}dx + \theta_f T\left(z\right) - C\left(\tilde{\lambda}_f,\theta_f\right)$$

Integral out, we have

$$\max_{\lambda_f,\theta_f} \tilde{\lambda}_f\theta_f \left(\frac{\mu z}{\mu-1}\max\left(\lambda_f,\frac{c}{z}\right)^{1-\mu} - c\max\left(\lambda_f,\frac{c}{z}\right)^{-\mu}\right) + \theta_f T\left(z\right) - C\left(\tilde{\lambda}_f,\theta_f\right) \tag{17}$$

We focus on the case where $\lambda_f \leq \frac{c}{z}$. So the objective function can be written as $\phi\left(z\right)\theta_f\tilde{\lambda}_f + \theta_f T\left(z\right) - C\left(\tilde{\lambda}_f,\theta_f\right)$, where $\phi\left(z\right) = \frac{1}{\mu-1}c^{1-\mu}z^{\mu}$ is an increasing function of z. Then the Focs are

$$FOC_{\tilde{\lambda}_f} = \phi\left(z\right)\theta_f - C_{\tilde{\lambda}} = 0$$

$$FOC_{\theta_f} = \phi\left(z\right)\tilde{\lambda}_f + T\left(z\right) - C_{\theta} = 0$$

The Hessian matrix is

$$SOC = \begin{bmatrix} -C_{\tilde{\lambda}\tilde{\lambda}} & \phi\left(z\right) \\ \phi\left(z\right) & -C_{\theta\theta} \end{bmatrix}$$

Given the assumptions, we have $C_{\tilde{\lambda}\tilde{\lambda}}C_{\theta\theta} > C_{\tilde{\lambda}\theta}^2 \geq \phi^2\left(z\right)$. So Hessian matrix is negative definite suggesting that the Focs are the sufficient conditions for the optimal solutions.

For the FOCs, take derivative with respect to T, we have

$$\begin{aligned}\phi d\theta_f &= C_{\tilde{\lambda}\theta} d\theta_f + C_{\tilde{\lambda}\tilde{\lambda}} d\tilde{\lambda}_f \\ dT &= C_{\theta\theta} d\theta_f - \phi d\lambda_f + C_{\tilde{\lambda}\theta} d\tilde{\lambda}_f\end{aligned}$$

Hence

$$d\tilde{\lambda}_f = -\frac{C_{\tilde{\lambda}\theta} - \phi}{C_{\tilde{\lambda}\tilde{\lambda}}} d\theta_f$$

$$dT = \left(C_{\theta\theta} - \frac{\left(C_{\tilde{\lambda}\theta} - \phi\right)^2}{C_{\tilde{\lambda}\tilde{\lambda}}} \right) d\theta_f$$

Under our assumptions, we can verify that $\frac{d\theta_f}{dT} > 0$ and $\frac{d\tilde{\lambda}_f}{dT} < 0$.

C.2 Proof of proposition 2

Let $\hat{x}$ denote the worst patent that will be sold to the z_H. Let x_0 be the patent that makes the surplus to be 0 ($S(z_H, x_0) = 0$) and x_1 be the patent that makes the surplus of z_L to be 0 ($S(z_L, x_1) = 0$).

First, we can verify that without the subsidy policy, for any $x \leq \frac{c}{z_H}$, the surplus $S(z_H, x)$ is 0 and if $x > \frac{c}{z_H}$ the surplus is $S(z_H, x)$ positive. Second, notice that if all z_H can purchase patents, then that is patents higher than $\lambda_i \left(\frac{\theta_i}{m(1-\sigma)G_H}\right)^{\frac{1}{\mu}}$. Hence without the subsidy, $\hat{x}_H = \max(\frac{c}{z_H}, \lambda_i \left(\frac{\theta_i}{m(1-\sigma)G_H}\right)^{\frac{1}{\mu}})$. All patents that higher than $\hat{x}_H$ will be sold to z_H. In this case, all the patents purchased by z_H will be renewed. The number of patents that are sold to z_H is $\theta_i \left(\frac{\lambda_i}{\hat{x}_H}\right)^{\mu}$.

Second, without the subsidy policy, we can verify that for any patents, $S(z_L, x) \leq 0$. Hence no patents will be sold to z_L without the subsidy.

Third, consider the case that subsidy is positive. We can verify that for any patents, the surplus $S(z_H, x)$ is always positive. So patents that higher than $\hat{x}'_H = \max\left(\lambda_i, \lambda_i \left(\frac{\theta_i}{m(1-\sigma)G_H}\right)^{\frac{1}{\mu}}\right)$

will be sold to z_H, where we use the prime to denote the cutoff after the policy is implemented. The max operator captures the possibility that if $\lambda_i \geq \lambda_i \left(\frac{\theta_i}{m(1-\sigma)G_H}\right)^{\frac{1}{\mu}}$, the number of patents from inventors is too small and all patents will be sold to z_H.

Meanwhile, we can verify that there exists a cutoff x_1, when $x \leq x_1$ the surplus $S(z_L, x)$ is positive, where $x_1 = c + T_L$ if $T_L < \frac{c}{z_L} - c$ or $x_1 = \frac{T_L}{1-z_L}$ if $T_L > \frac{c}{z_L} - c$. Notice that the patents sold to z_L must be lower than $\hat{x}_H^{'}$. And if all z_L have purchased patents, the cutoff of patents is $\lambda_i \left(1 - \frac{m(1-\sigma)G_L}{\theta_i}\right)^{-\frac{1}{\mu}}$. Hence we have patents lower than $\hat{x}_L' = \max(\lambda_{i,} \min[x_1, \hat{x}_H^{'}, \lambda_i \left(1 - \frac{m(1-\sigma)G_L}{\theta_i}\right)^{-\frac{1}{\mu}}])$ will be sold to the z_L.

C.3 Proof of proposition 3

Let We have the value of innovating for the inventor is

$$\begin{aligned}
& \theta_i \int W_i(x)\, dH_i(x) \\
= \; & \alpha\theta_i \int_{\hat{x}_H}^{+\infty} \max(z_H x - c, 0) dH_i(x) + \alpha\theta_i \int_{\lambda_i}^{\hat{x}_L} \max(z_L x - c, 0) dH_i(x) \\
& -\alpha\theta_i \int_{\hat{x}_H}^{+\infty} \max(x - c, 0)\, dH_i(x) - \alpha\theta_i \int_{\lambda_i}^{\hat{x}_L} \max(x - c, 0)\, dH_i(x) \\
& +\alpha\theta_i \int_{\hat{x}_H}^{+\infty} T_H dH(x) + \alpha\theta_i \int_{\lambda_i}^{\hat{x}_L} T_L dH(x) \\
& +\theta_i \int_{\lambda_i}^{+\infty} \max(x - c, 0)\, dH_i(x)
\end{aligned}$$

We focus on the case where $\lambda_i \leq c$ so that some patents will not be renewed.

Before the subsidy, patents sell to z_H will always be renewed and patents sell to z_L will never be renewed. And no patents will be sold to z_L. We have

$$\theta_i \int W_i(x)\, dH_i(x) = B(\hat{x}_H)\, \tilde{\lambda}_i \theta_i - C\left(\tilde{\lambda}_i, \theta_i\right)$$

where $B(\hat{x}_H)$ is an endogenous outcome, but the inventor takes it as given in the optimization. It is defined as

$$B(\hat{x}_H) = \alpha\left(\frac{\mu z_H}{\mu-1}\hat{x}_H^{1-\mu} - c\hat{x}_H^{-\mu}\right) - \alpha\left[\frac{\mu}{\mu-1}\max[\hat{x}_H, c]^{1-\mu} - c\max[\hat{x}_H, c]^{-\mu}\right] + \frac{1}{\mu-1}c^{1-\mu}$$

After the subsidy, some patents will be sold to z_L but they will not be renewed. We then have

$$\theta_i \int W_i(x)\, dH_i(x) = M(\hat{x}'_H, \hat{x}'_L)\,\tilde{\lambda}_i\theta_i + \alpha T_L\theta_i - C\left(\tilde{\lambda}_i, \theta_i\right)$$

where $M(\hat{x}'_H, \hat{x}'_L)$ is an endogenous outcome,

$$\begin{aligned}
M(\hat{x}'_H, \hat{x}'_L) \quad = \quad & \alpha\left(\frac{\mu z_H}{\mu-1}\max\left(\hat{x}'_H, \frac{c}{z_H}\right)^{1-\mu} - c\max\left(\hat{x}'_H, \frac{c}{z_H}\right)^{-\mu}\right) \\
& -\alpha\left[\frac{\mu}{\mu-1}\max[\hat{x}'_H, c]^{1-\mu} - c\max[\hat{x}'_H, c]^{-\mu}\right] \\
& -\alpha\left[\frac{\mu}{\mu-1}\max(c^{1-\mu} - \hat{x}_L'^{1-\mu}, 0) - c\max(c^{-\mu} - \hat{x}_L'^{-\mu}, 0)\right] \\
& +\alpha\left(T_H\hat{x}_H'^{-\mu} - T_L\hat{x}_L'^{-\mu}\right) + \frac{1}{\mu-1}c^{1-\mu}
\end{aligned}$$

The above equation has a similar form with the problem (17). So if T_L increases, there is a stronger incentive for the inventor to decrease the quality of patents because low quality patents can be directly sold to z_L.

D The planner problem

For any variable e, we denote e^o as the value under the optimal subsidy. First, the profit of the self-developed firm is

$$\begin{aligned}
\pi_{fH} &= \max_{\tilde{\lambda}^o_{fH},\theta^o_{fH}} \phi_H \tilde{\lambda}^o_{fH}\theta^o_{fH} + T_H\theta^o_{fH}\tilde{\lambda}^o_{fH}\bar{x}_s^{-\mu} - C_H\left(\tilde{\lambda}^o_{fH},\theta^o_{fH}\right) \\
\pi_{fL} &= \max_{\tilde{\lambda}^o_{fL},\theta^o_{fL}} T_L\theta^o_{fL}\tilde{\lambda}^o_{fL}\bar{x}_s^{-\mu} - C_L\left(\tilde{\lambda}^o_{fL},\theta^o_{fL}\right)
\end{aligned}$$

where $\tilde{\lambda}^o_{fH}\bar{x}_s^{-\mu}$ is the probability that the patent's quality is higher than $\bar{x}_s$ so that the firm can get the subsidy. We have

$$\tilde{\lambda}^o_{fH} = \tilde{\lambda}^o_{fL} = \frac{1+v_2}{1+v_{1H}+v_2}\tilde{\lambda}_{\max} \tag{18}$$

$$\phi_H\tilde{\lambda}^o_{fH} + T_H\tilde{\lambda}^o_{fH}\bar{x}_s^{-\mu} = C_{H,\theta}\left(\tilde{\lambda}^o_{fH},\theta^o_{fH}\right) \tag{19}$$

$$T_L\tilde{\lambda}^o_{fL}\bar{x}_s^{-\mu} = C_{L,\theta}\left(\tilde{\lambda}^o_{fL},\theta^o_{fL}\right) \tag{20}$$

Second we consider the inventor's problem. We focus on the case where $S\left(\hat{x}^o_H, z_H\right) > 0$ and $S\left(\hat{x}^o_L, z_L\right) > 0$. Then from the trade market clearing conditions, we have $\hat{x}^o_H = \lambda^o_i\left(\frac{\theta^o_i}{m(1-\sigma)G_H}\right)^{\frac{1}{\mu}}$ and $\hat{x}^o_L = \lambda^o_i\left(1-\frac{m(1-\sigma)G_L}{\theta^o_i}\right)^{-\frac{1}{\mu}}$ .Since $m\left(1-\sigma\right)$ is small, so we focus on the case that $\hat{x}^o_H > c$ and $\hat{x}^o_L < c$. Given the cutoff $\bar{x}_t$, patents that are above $\max\left(\hat{x}^o_H, \bar{x}_t\right)$ will be sold to z_H and patents between $[\bar{x}_s, \hat{x}^o_L]$ will be sold to z_L.

The profit of the inventor is

$$\pi_i = \max_{\tilde{\lambda}^o_i,\theta^o_i} M\left(\hat{x}^o_H,\bar{x}_t\right)\tilde{\lambda}^o_i\theta^o_i - C_i\left(\tilde{\lambda}^o_i,\theta^o_i\right)$$

where $M\left(\hat{x}^o_H,\bar{x}_t\right)$ is defined in a similar way as before.

Given the functional form of our assumption, we have

$$\tilde{\lambda}_i^o = \frac{1+v_2}{1+v_1+v_2}\tilde{\lambda}_{\max} \tag{21}$$

$$M\left(\hat{x}_H^o, , \hat{x}_L^o, \bar{x}_t\right) = C_{i,\theta_i^o}\left(\tilde{\lambda}_i^o, \theta_i^o\right) \tag{22}$$

The aggregate welfare is

$$\begin{aligned} U \quad = \quad & \left[\frac{\mu(z_H-1)}{\mu-1}\max(\hat{x}_H^o, \bar{x}_t)^{1-\mu} + \frac{1}{\mu-1}c^{1-\mu}\right]\tilde{\lambda}_i^o\theta_i^o - C\left(\tilde{\lambda}_i^o, \theta_i^o\right) \\ & + m\sigma G_H\left[\phi_H\tilde{\lambda}_{fH}^o\theta_{fH}^o - C\left(\tilde{\lambda}_{fH}^o, \theta_{fH}^o\right)\right] - m\sigma G_L C_L\left(\tilde{\lambda}_{fL}^o, \theta_{fL}^o\right) \\ & + \varepsilon\left[\frac{\mu}{\mu-1}\theta_i^o\tilde{\lambda}_i^{o\frac{1}{\mu}} + m\sigma G_H\frac{\mu}{\mu-1}\tilde{\lambda}_{fH}^{o\frac{1}{\mu}}\theta_{fH}^o + m\sigma G_L\frac{\mu}{\mu-1}\tilde{\lambda}_{fL}^{o\frac{1}{\mu}}\theta_{fL}^o\right] \end{aligned}$$

The first line is the profit created by the inventor's innovation, after removing the subsidy. The second line is the profit created by the firm's innovation. Notice that $z_L = 0$, so the net value is alwasy negative $-C_L\left(\tilde{\lambda}_{fL}^o, \theta_{fL}^o\right)$ after excluding the subsidy. The last line captures the knowledge spillover effect.

The planner maximizes the U by choosing $\bar{x}_s$ and $\bar{x}_t$, and subject to the constraints that all agents optimize their problems (equations (18-22)). Notice that we assume that the government has a deep pocket, so that there is the total subsidy does not show up as a constraint.

Let us first focus on the cutoff $\bar{x}_s$. The FOC of the planner is

$$\frac{\partial U}{\partial\theta_{fH}^o}\frac{\partial\theta_{fH}^o}{\partial\bar{x}_s} + \frac{\partial U}{\partial\theta_{fL}^o}\frac{\partial\theta_{fL}^o}{\partial\bar{x}_s} = 0$$

Using the firm's optimization conditions (equations (19-20)), we have

$$\begin{aligned}\frac{\partial U}{\partial \theta^o_{fH}} &= m\sigma G_H\left(\varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}^{o\frac{1}{\mu}}_{fH} - T_H\tilde{\lambda}^o_{fH}\bar{x}_s^{-\mu}\right)\\ \frac{\partial U}{\partial \theta^o_{fH}} &= m\sigma G_L\left(\varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}^{o\frac{1}{\mu}}_{fL} - T_L\tilde{\lambda}^o_{fL}\bar{x}_s^{-\mu}\right)\end{aligned}$$

Notice that the social value of θ^o_{fH} is $\phi_H\tilde{\lambda}^o_{fH} - C_{H,\theta}\left(\tilde{\lambda}^o_{fH},\theta^o_{fH}\right) + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}^{o\frac{1}{\mu}}_{fH}$ and the private value is $\phi_H\tilde{\lambda}^o_{fH} - C_{H,\theta}\left(\tilde{\lambda}^o_{fH},\theta^o_{fH}\right) + T_H\tilde{\lambda}^o_{fH}\bar{x}_s^{-\mu} = 0$. So $\frac{\partial U}{\partial \theta^o_{fH}}$ is the difference between the social value and private value of θ^o_{fH}, adjusted by the number of firms. Similarly, the social value of θ^o_{fL} is $\varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}^{o\frac{1}{\mu}}_{fL} - C_{L,\theta}\left(\tilde{\lambda}^o_{fL},\theta^o_{fL}\right)$, and the private value is $T_L\tilde{\lambda}^o_{fL}\bar{x}_s^{-\mu} - C_{L,\theta}\left(\tilde{\lambda}^o_{fL},\theta^o_{fL}\right) = 0$. So $\frac{\partial U}{\partial \theta^o_{fL}}$ is the $\theta^{o\prime}_{fL}s$ social value minus its private value. As we can see that as long as $\varepsilon > 0$, the social values will not be the same as the private values unless the planner subsidizes the innovations.

Solve $\bar{x}_s$ from the FOC, we have

$$\bar{x}_s = \left[\frac{\tilde{\lambda}^{o\left(1-\frac{1}{\mu}\right)}_{fH}\left(G_H\frac{T_H^2}{\theta^{o(v_2-1)}_{fH}} + G_L\frac{T_L^2}{\theta^{o(v_2-1)}_{fL}}\right)}{\varepsilon\frac{\mu}{\mu-1}\left(G_H\frac{T_H}{\theta^{o(v_2-1)}_{fH}} + G_L\frac{T_L}{\theta^{o(v_2-1)}_{fL}}\right)}\right]^{\frac{1}{\mu}} \tag{23}$$

So if ε increases, the planner chooses a smaller $\bar{x}_s$ to subsidize more firms. If $\varepsilon = 0$, $\bar{x}_s$ is infinity, so that the planner does not subsidize any firms.

Now consider the choice of $\bar{x}_t$. The social value of θ^o_i is $\frac{\mu(z_H-1)}{\mu-1}\max(\hat{x}^o_H,\bar{x}_t)^{1-\mu} + \mu(z_H - 1)\tilde{\lambda}^o_i\theta^o_i x^{o-\mu}_H\frac{\partial \hat{x}^o_H}{\partial \theta^o_i}\chi\left(\hat{x}^o_H \geq \bar{x}_t\right) + \frac{1}{\mu-1}c^{1-\mu} - C_{i,\theta} + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}^{o\frac{1}{\mu}}_i$, while the private value is $\alpha\frac{\mu(z_H-1)}{\mu-1}\max(\hat{x}^o_H,\bar{x}_t)^{1-\mu} + \alpha T_H\max(\hat{x}^o_H,\bar{x}_t)^{-\mu} + \alpha T_L\max(\bar{x}^{o-\mu}_t - \hat{x}^{o-\mu}_L, 0) + \frac{1}{\mu-1}c^{1-\mu} - C_{i,\theta} = 0$. There are three reasons why the social value of θ^o_i is different from the private value. First, $\alpha < 1$. So that the inventor only gets part of the benefit from the trade. So the classical hold-up problem will make the innovation incentive weaker in the decentralized economy. Second, when an individual inventor makes more innovation, it decreases the trade probability of other inventors.

But the inventor does not internalize this effect. So this channel may make the incentive of innovation stronger than the optimal case. Third, the knowldege spillover is not internalized. The planner chooses $\bar{x}_t$, and the FOC is

$$\frac{\partial U}{\partial \theta_i^o}\frac{\partial \theta_i^o}{\partial \bar{x}_t} = 0$$

where $\frac{\partial U}{\partial \theta_i^o}$ could be shown as the difference between the social value and the private value of θ_i^o.

$$\begin{aligned}\frac{\partial U}{\partial \theta_i^o} &= \frac{\mu(z_H-1)}{\mu-1}\max(\hat{x}_H^o,\bar{x}_t)^{1-\mu} + \mu(z_H-1)\tilde{\lambda}_i^o\theta_i^o x_H^{o-\mu}\frac{\partial \hat{x}_H^o}{\partial \theta_i^o}\chi\left(\hat{x}_H^o \geq \bar{x}_t\right) + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}} \\ &\quad - \left[\alpha\frac{\mu(z_H-1)}{\mu-1}\max(\hat{x}_H^o,\bar{x}_t)^{1-\mu} + \alpha T_H\max(\hat{x}_H^o,\bar{x}_t)^{-\mu} + \alpha T_L\max(\bar{x}_t^{-\mu} - \hat{x}_L^{o-\mu},0)\right]\end{aligned}$$

Since $\frac{\partial \theta_i^o}{\partial \bar{x}_t} > 0$, so we need to choose $\bar{x}_t$ such that $\frac{\partial U}{\partial \theta_i^o} = 0$.

There are three regions of $\bar{x}_t$. First, $\bar{x}_t < \hat{x}_L^o$. In this case,

$$\begin{aligned}\frac{\partial U}{\partial \theta_i^o} &= \left[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1} - \alpha T_H\right]\hat{x}_H^{o1-\mu} - \mu(z_H-1)\tilde{\lambda}_i^o\theta_i^o x_H^{o-\mu}\frac{\partial \hat{x}_H^o}{\partial \theta_i^o} + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}} - \alpha T_L\left(\bar{x}_t^{-\mu} - \hat{x}_L^{o-\mu}\right) \\ &= [(1-\alpha)\frac{\mu(z_H-1)}{\mu-1} - \alpha T_H - (z_H-1)\tilde{\lambda}_i^o]\hat{x}_H^{o1-\mu} + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}} - \alpha T_L\left(\bar{x}_t^{-\mu} - \hat{x}_L^{o-\mu}\right)\end{aligned}$$

Second, $\hat{x}_L^o < \bar{x}_t < \hat{x}_H^o$. Then

$$\begin{aligned}\frac{\partial U}{\partial \theta_i^o} &= \left[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1} - \alpha T_H\right]\hat{x}_H^{o1-\mu} - \mu(z_H-1)\tilde{\lambda}_i^o\theta_i^o\hat{x}_H^{o-\mu}\frac{\partial \hat{x}_H^o}{\partial \theta_i^o} + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}} \\ &= [(1-\alpha)\frac{\mu(z_H-1)}{\mu-1} - \alpha T_H - (z_H-1)\tilde{\lambda}_i^o]\hat{x}_H^{o1-\mu} + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}}\end{aligned}$$

Third, $\bar{x}_t \geq \hat{x}_H^o$. Then

$$\frac{\partial U}{\partial \theta_i^o} = \left[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1} - \alpha T_H\right]\bar{x}_t^{1-\mu} + \varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}}$$

If $[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1}-\alpha T_H-(z_H-1)\tilde{\lambda}_i^o]\hat{x}_H^{o1-\mu}+\varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}}>0$, so α is very small. Then in the second and third case $\frac{\partial U}{\partial\theta_i^o}>0$, or $\frac{\partial U}{\partial\theta_i^o}\frac{\partial\theta_i^o}{\partial\bar{x}_t}<0$. So $\bar{x}_t$ should be as low as possible in the second and third case. So we need to choose $\bar{x}_t<\hat{x}_L^o$ and let $\frac{\partial U}{\partial\theta_i^o}=0$.

$$\bar{x}_t=\left[\frac{\alpha T_L}{[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1}-\alpha T_H-(z_H-1)\tilde{\lambda}_i^o]\hat{x}_H^{o1-\mu}+\varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}}+\alpha T_L\hat{x}_L^{o-\mu}}\right]^{\frac{1}{\mu}}$$

If $[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1}-\alpha T_H-(z_H-1)\tilde{\lambda}_i^o]\hat{x}_H^{o1-\mu}+\varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}}<0$ but $\left[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1}-\alpha T_H\right]>0$, so α is in the middle. Then in the first and second case $\frac{\partial U}{\partial\theta_i^o}\frac{\partial\theta_i^o}{\partial\bar{x}_t}>0$. But in the third case, $\frac{\partial U}{\partial\theta_i^o}\frac{\partial\theta_i^o}{\partial\bar{x}_t}<0$. So we need to choose $\bar{x}_t=\hat{x}_H^o$.

If $[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1}-\alpha T_H-(z_H-1)\tilde{\lambda}_i^o]\hat{x}_H^{o1-\mu}+\varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}}<0$ but $\left[(1-\alpha)\frac{\mu(z_H-1)}{\mu-1}-\alpha T_H\right]<0$, we choose $\bar{x}_t>\hat{x}_H^o$ and let $\frac{\partial U}{\partial\theta_i^o}=0$.

$$\bar{x}_t=\left[\frac{\alpha T_H-(1-\alpha)\frac{\mu(z_H-1)}{\mu-1}}{\varepsilon\frac{\mu}{\mu-1}\tilde{\lambda}_i^{o\frac{1}{\mu}}}\right]^{\frac{1}{\mu-1}}$$