Types of Patents and Driving Forces behind the Patent Growth in China*

Zhiyuan Chen† Jie Zhang‡

November 17, 2018

Abstract

As a developing economy, China's unprecedented patenting surge is puzzling. We study China's patent surge and its driving forces using a novel and comprehensive merged dataset on patent applications filed by Chinese firms. We find that R&D investment, FDI, and patent subsidy have different effects on different types of patents. First, R&D investment has a positive and significant impact on patenting activities for all types of patents under different model specifications. Second, the stimulating effect of foreign direct investment on patent applications is only robust for utility model patents and design patents. Third, the patent subsidy only has a positive impact on design patents. The results imply that FDI and patent subsidy may disproportionately spur low-quality patents.

Keywords: patent types; R&D; foreign direct investment; patent subsidies

1 Introduction

The number of patents in China has been exploding in the past three decades. Since 2011, China has become the world's number one in filing patent applications. Breaking the patent counts into invention patents, utility models and designs, this extraordinary growth prevails.[1] According to the National Bureau of Statistics of China, applications for invention patents had increased from 25,236 in 2000 to 293,066 in 2010, with an average annual growth rate of 31.17%. Meanwhile, the utility model

*We thank Sushanta Mallick (the editor) and three anonymous referees for their constructive comments. We also would like to acknowledge seminar participants at the Special Issue Conference on "The Challenges of Managing and Modelling Innovation and Growth in China". Zhang gratefully acknowledges funding from MOE project of Key Research Institute of Humanities and Social Sciences at Universities (Grant no. 17JJD790022). All remaining errors are ours.

†Department of Economics, Pennsylvania State University. Email: zxc5109@psu.edu.

‡Corresponding author, Institute of China Economic Reform and Development, Renmin University of China. Email: zhangjie0402@ruc.edu.cn.

[1]China National Intellectual Property Administration (CNIPA) classifies the patents into three categories. See detailed explanation on this classification in footnote 2.

(design) patents applications had also risen steadily with a growth rate of 19.86% (24.73%). If different types of patents represent distinct forms of innovation, patent heterogeneity should be important for understanding the driving forces behind China's patent surge as well as its policy implications.

Firms make the patenting decision by analyzing its costs and benefits. In principle, given the supply of new ideas, anything that affects the costs or benefits of patenting can influence the patenting outcome. Different ideas represent various types of innovation. As pointed out by Nemlioglu and Mallick (2017), different types of innovation may benefit firms unequally. To distinguish ideas by their novelty and applicability, CNIPA classifies the patents into three categories–invention, utility models, and designs.[2] These three types of patents vary greatly in the length of examination period, protection period as well as requirements for being granted. These differences presumably affect the net benefits of patenting, which further influence firms' incentives for patenting.

Using a novel combined database of Chinese manufacturing firms, this paper aims to deepen the understanding on Chinese patents growth by explicitly considering different types of patents separately. Unlike existing studies, we show that the patent heterogeneity is important in analyzing the patent surge in China. We apply count data models to deal with the problem of over-dispersion and excessive zeros in the patent counts data. Our study documents that factors explaining the patenting growth vary across different types of patents. The empirical results robustly show that R&D investment is one of the most important explanatory factors for all types of patents, but the marginal effect of R&D differs for different types of patents. Contrasting with Hu and Jefferson (2009), we find that foreign direct investment (FDI) only helps explain the creation of utility model and design patents, but not invention patents. More interestingly, different from Li (2012), our study shows that patent subsidy only has positive impact on the patents applications for designs. These results suggest the non-innovation motives are important in explaining the patent applications by firms, but their importance depends on the type of patents. This also implies that certain policies targeted at promoting patenting activities may distort firms' incentives and induce low-quality patents.

Our study relates to several strands of literature. First, it is closely related to the studies on patent surge in China and other countries. To the best of our knowledge, Hu and Jefferson (2009) is the first study documents and analyzes China's patent surge. Using a firm-level dataset that contains invention patents statistics on large and medium sized industrial enterprises, they show that R&D only explains a fraction of the explosive growth of invention patents. They find that FDI, amendments to the patent law in 2000 and ownership reform have fostered Chinese firms to file more applications for invention patents. In addition to these factors, the stimulating effect of patent subsidy programs ini-

[2]In China's patent law, invention is referred to the new technical solution proposed for the product, method or related improvement; the utility model refers to a new technical solution suitable for practical use proposed for shape, construction or their combination. According to Article 22 of the Patent Law of the P.R.C.: any invention or utility model for which patent right may be granted must possess novelty, inventiveness and practical applicability. In comparison, the requirement for the approving of design patents is in Article 24 of the Patent Law of the P.R.C as "…… must not be identical with or similar to any design which, before the date of filing, has been publicly disclosed in publications in the country or abroad or has been publicly used in the country, and must not collide with any prior legal rights obtained by any other person."

tiated by Chinese provincial governments is also documented by Li (2012) and Dang and Motohashi (2015). A recent study by Hu, Zhang, and Zhao (2017) shows that R&D has become less important in explaining the patenting applications. They also document a weaker correlation between patents and labor productivity. Put together, these studies suggest that the patent growth in China is not only driven by the intensification of R&D but also by other non-innovation motives. Studies on the patent surge in U.S. and Japan indicate that the impact of strengthened IPR protection on patenting is limited. Kortum and Lerner (1999) find that the jump in U.S. patenting between 1985 and 1995 is mainly spurred by the shift in the management of research towards more applied activities but not by the seemingly pro-patent legislative changes in the 1980s. Similarly, Sakakibara and Branstetter (2001) examine the impact of the 1988 Japanese patent law reforms, and also find no evidence supporting that the expanding of patent protection increased the R&D spending or patents.

Second, this study is associated with the literature on the technological effects of FDI. FDI can influence domestic firms through positive agglomeration effects or negative competition effects (Aitken and Harrison, 1999). Since the China's open policy initiated in the 1980s, FDI has played an important role in stimulating China's economic growth. The technological spillovers from FDI, however, remains unclear. This is probably because institutional factors such as the protection of intellectual property rights affect the magnitude of FDI spillovers (Bournakis, Christopoulos, and Mallick, 2018). Using a provincial dataset from 1995 to 2000, Cheung and Ping (2004) find positive impact of FDI on the number of (all types of) domestic patent applications in China. Using panel data analysis on Chinese high-tech industries, Liu and Buck (2007) find the sources of foreign technology spillovers and absorptive ability jointly determine the R&D performance of domestic firms. Nevertheless, Chen, Wang, and Singh (2018) show the domestic private investment has become the dominant contributor to China's technological progress. They notice that the state-owned investment and FDI actually reduce the impact of domestic private investment on stimulating technological advancement. A more comprehensive evaluation of the FDI spillovers by Lu, Tao, and Zhu (2017) reveals a negative impact of horizontal FDI, i.e., FDI in the same industry, upon the productivity by Chinese domestic firms. They also find no significant impact of FDI on spurring new products. We also find that FDI has no significant impact on the filings of invention patents, suggesting that the technology spillovers from FDI is limited. Moreover, in our study FDI is found to have significant and positive effects on the patenting for utility models and designs. This implies that firms may employ the patenting for low-quality ideas as a strategic tool to preempt competition from foreign firms. Policies aiming to promote domestic technological progress through attracting FDI may have unintended consequences by inducing firms to produce low-quality patents.

Lastly, our study connects to the literature on the effectiveness of patent-related fees in screening patent quality. Patent fees are an essential element in the design of patent system. A large body of literature has discussed the use of fees as a policy tool to weed out low-quality patents (see Caillaud and Duchêne, 2011; De Rassenfosse and Jaffe, 2018; Gans, King, and Lampe, 2004; Schankerman and

Pakes, 1986; Scotchmer, 1999). Our study fits into this strand of literature by focusing on reductions in patenting application fees and examination fees caused by provincial innovation subsidy programs. The empirical results suggest that the decrease in patenting fees induces more design patents. As we mentioned earlier, design patents are of the lowest quality among all types of patents. In this sense, the result suggests that maintaining certain level of patent fees is necessary for screening out low-quality patents. Moreover, this may also reflect that the impact of patenting subsidy on stimulating the firm's innovation is limited. This is probably because patenting fees are small relative to the expected return from patents granted for inventions and utility models.

This paper contributes to the existing literature in several aspects. First, this paper analyzes the patent surge in China by explicitly considering three types of patents separately. We evaluate a set of factors that may affect the patenting outcome for different types of patents in China. This enables us to detect the potentially different driving forces for different categories of patent that represent different forms of innovation, thus providing a more complete explanation of Chinese patents growth. Second, this study also has important implications for innovation policies. We find that R&D investment, FDI, and patent subsidy play different roles in spurring different types of patents. For example, if patent subsidy is only effective for stimulating low-quality patents, subsidizing on the patenting fees may cause a surge in low-quality patents that harm innovation incentives (Barton, 2000; De Rassenfosse and Jaffe, 2018). Finally, this study also has general implications for researches using patents to measure innovation activities. In addition to R&D investment, other factors may also affect the firm's patenting choice. In this case, using patenting as the measure of innovation regardless of the institutional setting can be misleading. Moreover, innovation can take place in different forms. Different innovation outcome have different market value, using aggregate measures such as R&D investment or the total number of patents disregard the quality of innovation.

The rest of this paper is organized as follows. We introduce the data used in this paper in Section 2. In Section 3, we display the descriptive statistics to motivate the formal econometric analysis on the driving forces behind patents. Section 4 shows the results in the order of the sophistication of the econometric models. Section 5 deals with the potential endogeneity problem. In Section 6 we conclude by discussing the empirical results and relevant policy implications.

2 Data

2.1 Data sources

This paper uses three databases. The first is a database of Chinese manufacturing firms from 2001 to 2007 compiled by China's National Bureau of Statistics (NBS). This dataset is widely used in economic studies focusing on China (see, for example, Hsieh and Klenow, 2009; Song, Storesletten, and Zilibotti, 2011; Chen, Zhang, and Zheng, 2017). It includes SOEs (State Owned Enterprises) and non-SOEs with annual sales no less than five million *Renminbi* (equivalent to about $700,000). These firms account for 98% of China's total manufacturing exports. The dataset includes more than 100 financial variables listed in the major accounting sheets of all these firms. In particular, it contains information on a firm's annual R&D expenditures.[3]

This study also uses a patents database provided by the CNIPA. It contains information on patent applications that are submitted by firms in mainland China. For each patent the database has information on its type (invention, utility model, or design), owner, application time, certification time, agent of application, abstract, location, and expiration time. This information essentially allows the researcher to track the entire life of patents from 1985 to 2012. But it should be noted that we have no information on patent citations, which makes it difficult to measure the patents quality directly.[4] Different from existing literature, we deal with different patents separately by assuming independence between different types of patents.[5] This allows us to identify different driving forces behind the surge in different kinds of patents. For the purpose of this study, we merge this database with data on Chinese manufacturing firms. The merged dataset contains the information of aforementioned two datasets. In Panel A of Table 1, we report the ratio of the number of patents applications of the merged dataset to the total number patent applications in mainland China. As we can see, the percentage of the total number of patents that are merged to the dataset is increasing for invention patents and utility model patents, decreasing for design patents. This implies that firms have played a more and more important role in producing high-quality patents in our sample, which adds to the importance of this study in understanding the long-term economic growth driven by Chinese firms. Since our dataset only covers around 10% of the total patent applications in China during 2001 to 2007, we have to be conservative should the results be generalized to the overall analysis of China's patent surge. Other entities such as research institutes and universities have also contributed to the patent growth in China (Li, 2012). Another concern is on the efficiency of the matching between two datasets. In the last row of Panel A in Table 1, we show the percentage of the total number of invention patents in the merged dataset to the figure in the China Statistical Yearbook on Science and Technology. We find this ratio varies across years, with 55.57% in 2007 and 96.35% in 2003. We tend to believe that the merged dataset is representative enough for the purpose of this study.

Table 1 here

The last dataset this study employs is the information on provincial patent subsidies from 2001-2007. This database is constructed from official documents released on the websites of the provincial intellectual property offices. For each type of patents, the patent subsidy policy is classified into five

[3]R&D expenditures are only available for observations no later than 2001, which restricts the time span of the study to be from 2001 to 2007.

[4]Dang and Motohashi (2015) propose to use the 'knowledge width' of each patent as a measure of its quality. This methodology uses the number of nouns in claims to quantify the claim scope; a wider scope of claim represents better patent quality.

[5]Our results are robust if we assume certain correlation between the equation of different types of patents.

categories based on various fees and statuses related to the application process, which include reductions in application fees, examination fees, agency costs, and annual fees, as well as grant-contingent rewards after the approving of the application. For each subsidy variable, the outcome can be defined according to one of the three possible states: no subsidy, partial subsidy, and complete subsidy; they are exclusive to each other. Dang and Motohashi (2015) measure the intensity of subsidy by assigning a larger value to complete subsidy than to partial subsidy. To reduce the measurement error caused by almost arbitrarily assigning values for the policy variables, we define the subsidy variable as a dummy variable which is equal to 1 for either partial subsidy or complete subsidy happens and 0 otherwise. This approach also rules out the potential differences between different sub-classes of patent subsidies imposed on different stages of patenting.[6] This full database is reported in Table A1 in the appendix. The final database includes the starting year of the implementation of patent subsidy in China mainland provinces from 2001 to 2007.

2.2 Data features

In this subsection, we present some descriptive statistics to motivate the formal econometric analysis in subsequent sections. In general, the preliminary description of the dataset stresses the importance of patent types in investigating the driving forces behind China's patents surge, especially for policies aimed at improving the innovative ability of Chinese firms.

2.2.1 Extensive and intensive margins of patent applications

How active are Chinese firms actually in applying for patents? Not surprisingly, a great majority of Chinese firms have never submitted any patent applications. A large fraction of zero-patent counts is an important feature of our data; we will take this into consideration when specifying the econometric models. The average percentage of firms filing patents applications are 1.11%, 1.25%, 1.99% for invention, design, and utility model patents, respectively. It is worth noting that the number of firms that produce patents is much lower than those have positive R&D expenditures. In our data, 11% of the entire observations are actively investing in R&D, which implies that R&D will not be fully transformed into patent ideas. This confirms that patenting is just one mechanism through which firms protect their profits due to innovation (Cohen, Nelson, and Walsh, 2000).

Although on average Chinese firms are not very active in applying for patents, the percentage of firms applying for invention and utility model patents is in a steady growing trend during our sample period. In Panel A of Figure 1, we plot the trends of patent applications. It shows clearly that the percentage of firms who are applicants for invention patents had reached 1.6% in 2007, almost 3 times larger than it was in 2001. Applications for utility model patents displayed a similar trend; the percentage increased from 1.45% in 2001 to 2.48% in 2007. In contrast, the percentage for active

[6] These include reductions in the fees associated with the application, examination, granting, and maintenance of patents.

design patents applicants is relatively stable, hovering around 1.2%. In 2006 the percentage of firms applying for invention patents had exceeded that number for design patents. This tells us that firms had become more active in generating invention and utility model patents. As we have mentioned, invention patents and utility model patents are of higher quality than design patents. The evolving patterns of extensive margin of patent applications show that more and more Chinese firms are applying for high-quality patents. Hu, Zhang, and Zhao (2017) also find that most of China's patenting growth is due to the expansion at the extensive margin during 2007 and 2011.

Figure 1 here

We also note that many firms apply for more than one type of patents. To provide a more complete picture of the change in Chinese firm's patent applications, we group the applicants into three cases–single-type patent applicants, two-type patent applicants, and three-type patent applicants (see Panel B of Figure 1). It is clear to see a steady growing pattern for all three cases. In particular, the percentage for two-type patent applicants has increased the most, three-type patent applicants the least. This indicates that firms are expanding the variety of patents.

After exploring the extensive margin of the patent application, we turn to describe the characteristics of average patent applications in the dataset. In Panel B of Table 1, we show the evolution of patent intensity defined as the average number of patent applications per firm between 2001 and 2007. As it shows, the average number of invention patents had increased the most, climbing from .016 up to .117. In 2002 and 2004, the increase is around twofold. In comparison, the trend of utility model patents and design patents are smoother. The average number of design patents even exhibited a decreasing trend during 2004-2006, though it bounded back to .097 in 2007.

2.2.2 R&D and patent applications

R&D measures the innovation motives for patenting. The significant positive relationship between R&D and patents has been well documented in many studies (Griliches, 1979, 1981; Hausman, Hall, and Griliches, 1984; Hu and Jefferson, 2009; Hu, Zhang, and Zhao, 2017). Motivated by this literature, we first check the simple correlation between R&D activities and the firm's decision to apply for patents for each type of patents, separately. Panel C of Table 1 shows how the contemporaneous and lagged R&D expenditures are associated with the patent application. Interestingly, many firm observations are found to have positive patent applications even in the absence of R&D investment; the pattern is quite similar for either present or lagged R&D activities. We find some weak evidence suggesting that firms undertaking R&D investment file more patent applications for inventions and utility models. In contrast, we find no evidence supporting that more innovative firms file more patent applications for new designs. At least the preliminary statistics show that non-R&D firms file more design patent applications than R&D-active firms.

Our simple statistics have shown that R&D have heterogeneous effects on the firm's behavior of filing different types of patents. In addition, non-R&D incentives may play a role in explaining the filings of patents applications for designs. Disregarding the substantial heterogeneity when investigating the driving forces of the firm's patent application will bias the estimates and generate misleading results.

3 Empirical strategy

In this section, we employ formal econometric methods to analyze the driving forces behind the surge in different types of patents. To save space, we only present two variations of count data models to deal with over-dispersion and excessive zeros in the data.

3.1 Over-dispersion and negative binomial regression models

3.1.1 Over-dispersion in the data

We use N_{it} to denote the number of patent applications, the basic specification of Poisson model is to parameterize the counts of patents as a Poisson distribution with mean λ_{it} that is associated with certain firm characteristics:

$$\Pr\left(N_{it} = n_{it} | \boldsymbol{X}_{it}\right) = \frac{e^{-\lambda_{it}} \lambda_{it}^{n_{it}}}{n_{it}!},$$
$$\lambda_{it} = \exp\left(\mathbf{X}_{it}' \boldsymbol{\beta}\right)$$

where $\mathbf{X}_{it}'$ includes interested explanatory variables. In particular, it can be written as

$$\mathbf{X}_{it}' = \left(\log\left(RD_{it}\right), \log\left(RD_{it-1}\right), \mathbf{Z}_{it}'\right)$$

where $\log\left(RD_{it}\right)$ is the log of R&D expenditures, and $\log\left(RD_{it-1}\right)$ is the lagged R&D expenditures. $\boldsymbol{Z}_{it}$ is a vector including other important factors documented in the literature, such as foreign direct investment (fdi_{it}) and\or patent subsidy ($psub_{it}$). To account for the industrial specific effects, we also include a industrial dummy variables that is equal to one when the firm belongs to a high-tech industry.[7] Considering that time effects may affect the growth of patents, we also control for the fixed effects. The specification of Poisson model implicitly imposes the conditional mean and variance of N_{it} are the same. When the conditional variance is greater than the conditional mean, the data are over-dispersed. To detect whether there is over-dispersion in the data, we allow the variance-mean ratio to be any positive constant:

[7] Alternatively, we have tried to include a full set of industry dummies to control for potential industry fixed effects; the results remain robust.

$$\text{Var}\left(N_{it}|\boldsymbol{X}_{it}\right) = \sigma^2\mathbb{E}\left(N_{it}|\boldsymbol{X}_{it}\right) = \sigma^2\lambda_{it} \tag{1}$$

Then σ^2 can be estimated using the QMLEs of $\boldsymbol{\beta}$ by considering the sample analog. We estimate the variance-mean ratio for the Poisson models. For all models, the estimation results show that $\hat{\sigma}$'s are greater than one, strongly suggesting over-dispersion in the data. Therefore we use alternative specifications to suit our data better.[8]

3.1.2 Negative binomial mean-dispersion model

Following the existing literature on count data models, we consider two approaches to deal with the over-dispersion in the data. One is the mean-dispersion model with a common parameter, the other is a model with a parameterized distribution for the unobserved heterogeneity. Below we only discuss briefly about these two methods. Details about these econometric methods can be found in Cameron and Trivedi (2013).

Mean-dispersion model with common parameter One way of constructing the negative binomial mean-dispersion model is to introduce unobserved heterogeneity in a Poisson model.[9] Let η_{it} be the unobserved error term and assume that $\exp\left(\eta_{it}\right)$ follow a gamma distribution with parameters $\left(1/\alpha, \alpha\right)$. Thus $\mathbb{E}\left(\exp\left(\eta_{it}\right)\right) = 1$, $\text{Var}\left(\eta_{it}\right) = \alpha$. We further assume that $\exp\left(\eta_{it}\right)$ is independent of $\boldsymbol{X}_{it}$.[10] It can be shown that the conditional distribution of N_{it} on $\boldsymbol{X}_{it}$ is negative binomial, with conditional mean and variance as follows.[11]

$$\mathbb{E}\left(N_{it}|\boldsymbol{X}_{it}\right) = \lambda_{it} \tag{2}$$

$$\text{Var}\left(N_{it}|\boldsymbol{X}_{it}\right) = \lambda_{it} + \alpha\lambda_{it}^2 \tag{3}$$

α is called the parameter of over-dispersion; the larger α is, the more over-dispersed the data are. In particular, when $\alpha = 0$, the negative binomial specification degenerates to be the Poisson model. In this sense, the negative binomial model generalizes the Poisson model to capture the over-dispersion in the data. As is pointed out, under (2), for any fixed positive value for α, the coefficient estimates of $\boldsymbol{\beta}$ by maximizing the associated log likelihood function $\mathcal{L}\left(\boldsymbol{\beta}, \alpha\right)$ are consistent (Wooldridge, 2010).

Mean-dispersion model with parameterized α Although the aforementioned negative binomial mean-dispersion model tackles the over-dispersion in the data to some extent, the assumption that α is identical to all observations is too restrictive. To relax this restriction, we consider parameterizing

[8] To save space, we do not present the detailed results. The full estimation results are available upon request.

[9] The negative binomial mean-dispersion model is also known as the NegBin Iwemodel (Cameron and Trivedi, 2013)

[10] We will return to discuss this assumption when we consider the endogeneity issue of our models.

[11] See Cameron and Trivedi (2013) for details about the derivation of the negative binomial model.

α as follows

$$\ln(\alpha_{it}) = \gamma_0 + \boldsymbol{W}'_{it}\boldsymbol{\theta} + f_o + f_t \tag{4}$$

where $\boldsymbol{W}'_{it}$ is a vector containing *age*, *size*, and *htech*. f_o represent the coefficients of ownership dummies; f_t is the coefficient for year dummies. γ_0, $\boldsymbol{\theta}$, f_o, and f_t are parameters to be estimated along with the other model coefficients.

3.2 Excessive zeros of patent applications

In the data section, we have shown that most (around 98%) of the firms do not file any patent applications. This poses challenges to the assumption that the outcome of patents follows a Poisson distribution. To fit the data better, we need to model the event of whether a firm creates patents in addition to the Poisson distribution. Lambert (1992) develops the Zero Inflated Poisson (ZIP) model to deal with this situation. Let p_{it} be the probability that a firm refuses to apply for any patents, then $(1 - p_{it})$ becomes the propensity of applying for some patents. To model the discrete choice to patent or not, we specify following logit model:

$$p_{it} = F\left(\mathbf{W}'_{it}\gamma\right) = \frac{1}{1 + \exp\left(-\mathbf{W}'_{it}\gamma\right)} \tag{5}$$

Then the log likelihood function for this specification can be written as:

$$\begin{aligned} L\left(\gamma, \boldsymbol{\beta}; n_{it}, \mathbf{X}_{it}, \mathbf{W}_{it}\right) = & \sum_{n_{it}=0} \ln\left\{F\left(\mathbf{W}'_{it}\gamma\right) + \left[1 - F\left(\mathbf{W}'_{it}\gamma\right)\right] \times \left[\exp\left(-\exp\left(\mathbf{X}'_{it}\boldsymbol{\beta}\right)\right)\right]\right\} \\ & + \sum_{n_{it}>0} \left\{\ln\left[1 - F\left(\mathbf{W}'_{it}\gamma\right)\right] - \exp\left(\mathbf{X}'_{it}\boldsymbol{\beta}\right) + n_{it}\mathbf{X}'_{it}\boldsymbol{\beta} - \ln\left(n_{it}!\right)\right\} \end{aligned} \tag{6}$$

The estimate of $(\gamma, \boldsymbol{\beta})$ is obtained by maximizing the above likelihood function (6). To account for the over-dispersion problem, we also add an unobserved component to λ_{it} and estimate a Zero Inflated Negative Binomial (ZINB) model.

4 Estimation results

The estimation results of the negative binomial mean-dispersion model are reported in Table 2. First, in each group of the models, the estimation results of α indicate strong over-dispersion in the data. Considering this, the negative binomial model, which takes the over-dispersion into account, fits our data better. In the results of negative binomial model, the patent subsidy is more effective for utility model patents and design patents. Also, private firms and foreign firms are found to file more patent applications than state-owned firms for all types of patents.

Table 2 here

Note that the productivity of R&D in creating patents differs across types of patents. The invention patents have the largest R&D-patent elasticity both for the current and lagged R&D, while the design patents display the smallest R&D-patent elasticity. Moreover, the coefficient of lagged R&D is smaller than that of the present R&D for all types of patents. However, for utility models and designs, R&D expenditures become less important for the filings of patents when we control for FDI and patent subsidy. Furthermore, FDI is more effective in stimulating the patenting applications for designs than inventions and utility models. Note that these differences found for various types of patents would disappear if we pool all patents together or only consider a single type of patents. Especially, the estimation results would be much less informative since the potential heterogeneous effects are averaged out when we pool all of the three types of patents together, which are shown in Column (1) in Table 2.

In Table 3 we display the estimation results for the mean-dispersion model with a parameterized α. As we can see from the results, the coefficients of explanatory variables display a pattern similar to results presented in Table 2. This shows that previous results are robust to the alternative parameterization of α. Also note that for the equation of $\ln(\alpha)$, the coefficients of *size*, *age*, and *htech* are all negative, implying that big firms, old firms, and firms in high-tech industries display smaller over-dispersion in the data.

Table 3 here

In Table 4, we present the results of ZIP and ZINB. For different types of patents, we estimate two different models by including different covariates into $\mathbf{W}_{it}$. In subgroup a, we include patent subsidies, firm size, firm age, ownership dummies, and a constant, while in subgroup b FDI and a full set of year dummies are added. Under our specification, because most of the observations are of zero patent applications, the inflate part of the model captures more about the extensive margin of the patents application. The part of Poisson process is associated more with the intensive margin of the patent application. For any variables included in $\mathbf{X}_{it}$, we say there is a strong evidence showing that it explains the patents outcome when its coefficient is significantly positive in the part of negative binomial model.

Table 4 here

There is still a lack of strong evidence showing that patent subsidy stimulates firms to file more invention patents. But the coefficients of $\log(RD_{it})$ and $\log(RD_{it-1})$ are positive and significant at

1% significance level for the group of invention patents. More importantly, these coefficients are larger than those in the groups of utility model patents or design patents, which implies R&D plays a more important role in explaining the invention patents. Last but not the least, the effect of FDI on invention patents is positive. This could be either the impact of foreign competition or knowledge spillovers (Aitken and Harrison, 1999; Lu, Tao, and Zhu, 2017). Overall, these results suggest that FDI is not a significant factor in explaining the patenting outcome when conditioning on the firm's R&D investment.

We can see from the middle columns of Table 4 that both patent subsidy and FDI play significant roles in driving up the number of applications for utility model patents. In all estimations, the coefficient of the dummy variable for private firms are significant and positive for the negative binomial part, and negative for the inflate part. This implies private firms are filing more patents for all of the three types of patents. In contrast, there is only evidence showing that foreign firms are filing more design patents.

5 Endogeneity issue

Our empirical results in previous section display some interesting patterns that are consistent with existing studies by Hu and Jefferson (2009) and Li (2012). However, we should be cautious because of the endogeneity issue caused by unobserved idiosyncratic characteristics. In this section, we try to use panel data methods to deal with this concern. We regard the estimation results as a more convincing interpretation of our dataset.

In the specification of negative binomial model, we allow the variance-to-mean ratio to differ across firms and across time by imposing a specific form to the Poisson parameter. However, a shortcoming of the negative binomial specification is it assumes that the unobserved firm-specific heterogeneity is independent of the explanatory variables. To deal with this problem, we follow Hausman, Hall, and Griliches (1984) (HHG hereafter) to estimate a fixed-effects negative binomial model. This model allows for arbitrary dependence between the unobserved idiosyncratic characteristics and the explanatory variables while allowing for over-dispersion in the data.[12] An alternative specification of the error term is the random-effects negative binomial model. We omit the details of its specification. Instead, we show results of the Hausman specification test that favor the fixed-effects model. To save space, we only report the estimation results of fixed effects negative binomial model in Table 5.

Table 5 here

One noteworthy result in Table 5 is a great shrinkage in the sample size. Compared to the original

[12]We also tried to employ the results of fixed effects Poisson model. The results, however, show that fixed effects Poisson model provides a poor fitting to the data with the Hausman testing statistic being negative for most of the groups.

sample size in the pooled regression, most of the observations are deleted due to the all-zero outcomes for the dependent variable. We also report the number of observations deleted because of single observation over the sample period. This great loss of information reminds us to be cautious when interpreting the results.

In Table 5, the coefficients of $\log{(RD_{it})}$ and $\log{(RD_{it-1})}$ are much smaller than estimates reported in Table 3 and Table 4. Recall that we have shown that there is a substantial R&D investment gap between firms with patent applications and those of no patent application. Since the panel data model drops most of the data with zero patent applications, the remaining dataset contains firms that invest relatively more in R&D. As a result, the variations in patents and R&D both become smaller. The coefficient estimates show that the drop in the variation of patents is more significant compared to that in the variation of investment in R&D. For the coefficient of FDI, we find it only drives the filings of utility model patents and design patents. When we look at the coefficient estimates for patent subsidy, the coefficient is only significantly positive for design patents.

6 Conclusion

R&D has a long-lasting effect on firm performance (Bournakis and Mallick, 2018). Our empirical results robustly show that R&D has a positive impact on patenting. This is consistent with the findings in studies investigating the R&D-patents relationship (Griliches, 1981; Hausman, Hall, and Griliches, 1984; Hall, Griliches, and Hausman, 1986). But a simple calculation can show that R&D growth is not sufficient to explain the patenting growth in China. According to the OECD database, China's R&D expenditures had a 256.27% (215.87%) increase from 1999 to 2006 (2000 to 2007). Given the patent-R&D elasticity reported in Table 5, the predicted patenting growth rate should be 5.64% for invention patents, 7.57% for utility model patents, and 9.24% for design patents. In contrast, the number of invention, utility model, and design patents have a 896.47% , 333.41%, and 160.05% increase in the sample period, separately. This reminds us to consider other factors in order to fully explain the patenting growth in China.

To some extent, the empirical results question the perception that foreign firms have stimulated Chinese firms to apply for more inventive patents. Conditional on the firm's R&D investment, FDI has no significant impact on the patenting for inventions. However, FDI has positive and significant effects on utility model and design patents. Note that the coefficient of FDI is a mixture of spillover effects and competition effects. This result implies that the spillover effect is restricted to low-quality ideas. With the reduced costs of innovation, the competition effect from FDI further encourages Chinese firms to employ relatively low-quality patents to take advantage of some loopholes in the the patent law to compete with foreign firms (Hu and Jefferson, 2009). Lu, Tao, and Zhu (2017) also find no beneficial spillovers from FDI in China.

We only find that the patent subsidy has positive and significant effects on the patenting of design patents. We attribute this finding to two main reasons. First, note that design patents are of the lowest quality. Generating a design patent application is of costs lower than an invention patent or utility model patent. Firms make patenting decision by comparing its expected payoffs and costs. If the reductions in the patenting fees are negligible compared to the benefits, the firm's patenting decision will not change. As a result, patenting subsidy will increase the low-quality patent applications disproportionately (De Rassenfosse and Jaffe, 2018). Second, Li (2012) argues that patent subsidies are not only effective in encouraging more individuals and universities to apply for invention patents but also are inducing firms to file more applications for invention patents. According to our results, we expect this argument only works for individuals and small firms which are usually more financially constrained than large firms and research institutions. Overall, patent subsidy has only stimulated the creation of low-quality patents (the design patents). This is consistent with the findings by Dang and Motohashi (2015). We also note that the coefficient of firm size is positive and significant for all the three different types of patents. This implies that larger firms are more likely to apply for patents. There is no significant correlation between firm age and the creation of patents; implying that patenting is neutral to firm age.

The empirical results have several implications for policies aiming to promote innovation in developing countries. First, development policies using FDI as the key driver of technological progress need to be reconsidered. FDI may play an important role in spreading relatively low-quality ideas, but relying on FDI to move up to the technological frontier is much less promising. The cutting-edge technology can only be developed through indigenous R&D. Second, patent subsidies increase the low-quality patents disproportionately by decreasing the patenting fees. The surge in low-quality patents may cause the fragmentation of intellectual property rights. Ultimately, the fragmentation will significantly raise the costs of using knowledge and may discourage R&D investment (Heller and Eisenberg, 1998). In addition, the surge in patenting applications may cause the patents examiners to spend less time on each patent and make more mistakes in granting patent rights. This can also lead to low-quality patents. To guarantee the quality standards, reductions in the patent fees should be combined with improvement in governance of patent offices as well as a supply of more professional patent examiners.

This study stresses the importance of patent quality in understanding the patent surge in China. As indicated by the empirical findings, R&D investment is more important in explaining high-quality patents, while FDI and patent subsidy stimulate the filings of patents of lower quality. It is the future work that we aim to reconcile these findings in a coherent theoretical framework.

References

Aitken, Brian J. and Ann E. Harrison. 1999. "Do domestic firms benefit from direct foreign investment? Evidence from Venezuela." *The American Economic Review* 89 (3):605–618.

Barton, John H. 2000. "Reforming the patent system." *Science* 287 (5460):1933–1934.

Bournakis, Ioannis, Dimitris Christopoulos, and Sushanta Mallick. 2018. "Knowledge spillovers and output per worker: An industry-level analysis for OECD countries." *Economic Inquiry* 56 (2):1028–1046.

Bournakis, Ioannis and Sushanta Mallick. 2018. "TFP estimation at firm level: The fiscal aspect of productivity convergence in the UK." *Economic Modelling* 70:579–590.

Caillaud, Bernard and Anne Duchêne. 2011. "Patent Office in innovation policy: Nobody's perfect." *International Journal of Industrial Organization* 29 (2):242–252.

Cameron, A. Colin and Pravin K. Trivedi. 2013. *Regression analysis of count data*, vol. 53. Cambridge University Press.

Chen, Hong, Xi Wang, and Baljeet Singh. 2018. "Can private domestic investment lead Chinese technological progress?" *Economic Modelling* 70:186–193.

Chen, Zhiyuan, Jie Zhang, and Wenping Zheng. 2017. "Import and innovation: Evidence from Chinese firms." *European Economic Review* 94:205–220.

Cheung, Kui-yin and Lin Ping. 2004. "Spillover effects of FDI on innovation in China: Evidence from the provincial data." *China Economic Review* 15 (1):25–44.

Cohen, Wesley M, Richard R Nelson, and John P Walsh. 2000. "Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not)." Working Paper 7552, National Bureau of Economic Research.

Dang, Jianwei and Kazuyuki Motohashi. 2015. "Patent statistics: A good indicator for innovation in China? Patent subsidy program impacts on patent quality." *China Economic Review* 35:137–155.

De Rassenfosse, Gaétan and Adam B Jaffe. 2018. "Are patent fees effective at weeding out low-quality patents?" *Journal of Economics & Management Strategy* 27 (1):134–148.

Gans, Joshua S, Stephen P King, and Ryan Lampe. 2004. "Patent renewal fees and self-funding patent offices." *Topics in Theoretical Economics* 4 (1):1–15.

Griliches, Zvi. 1979. "Issues in assessing the contribution of research and development to productivity growth." *The Bell Journal of Economics* 10 (1):92–116.

———. 1981. "Market value, R&D, and patents." *Economics Letters* 7 (2):183–187.

Hall, Bronwyn H, Zvi Griliches, and Jerry A Hausman. 1986. "Patents and R and D: Is there a lag?" *International Economic Review* 27 (2):265–283.

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches. 1984. "Econometric models for count data with an application to the patents-R&D relationship." *Econometrica* 52 (4):909–938.

Heller, Michael A and Rebecca S Eisenberg. 1998. "Can patents deter innovation? The anticommons in biomedical research." *Science* 280 (5364):698–701.

Hsieh, Chang-Tai and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4):1403–1448.

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

Hu, Albert GZ, Peng Zhang, and Lijing Zhao. 2017. "China as number one? Evidence from China's most recent patenting surge." *Journal of Development Economics* 124:107–119.

Kortum, Samuel and Josh Lerner. 1999. "What is behind the recent surge in patenting?" *Research Policy* 28 (1):1–22.

Lambert, Diane. 1992. "Zero-inflated Poisson regression, with an application to defects in manufacturing." *Technometrics* 34 (1):1–14.

Li, Xibao. 2012. "Behind the recent surge of Chinese patenting: An institutional view." *Research Policy* 41 (1):236–249.

Liu, Xiaohui and Trevor Buck. 2007. "Innovation performance and channels for international technology spillovers: Evidence from Chinese high-tech industries." *Research Policy* 36 (3):355–366.

Lu, Yi, Zhigang Tao, and Lianming Zhu. 2017. "Identifying FDI spillovers." *Journal of International Economics* 107:75–90.

Nemlioglu, Ilayda and Sushanta K. Mallick. 2017. "Do managerial practices matter in innovation and firm performance relations? New evidence from the UK." *European Financial Management* 23:1016–1061.

Sakakibara, Mariko and Lee Branstetter. 2001. "Do stronger patents induce more innovation? Evidence from the 1988 Japanese Patent Law reforms." *RAND Journal of Economics* 32 (1):77–100.

Schankerman, Mark and Ariel Pakes. 1986. "Estimates of the value of patent rights in European countries during the post-1950 period." *The Economic Journal* 96 (384):1052–1076.

Scotchmer, Suzanne. 1999. "On the optimality of the patent renewal system." *The RAND Journal of Economics* 30 (2):181–196.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing like China." *American Economic Review* 101 (1):196–233.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*. MIT press.

Tables and figures (in the order of presence)

Table 1 Summary statistics of patents in merged dataset

Panel A: Number and percentage of patents in the merged dataset

year	2001	2002	2003	2004	2005	2006	2007
invention	1982	4462	5333	7993	10100	17033	19750
	6.60%	11.21%	9.39%	12.15%	10.80%	13.93%	12.90%
utility model	4202	5649	7496	7798	10720	15324	18212
	5.30%	6.13%	6.95%	6.99%	7.76%	9.58%	10.12%
design	6316	8838	9131	10326	12665	15393	16425
	11.19%	12.01%	10.54%	10.17%	8.35%	8.19%	6.48%
matching efficiency	57.10%	81.26%	96.35%	87.20%	57.58%	67.33%	55.67%

Panel B: Trend of average number of patent applications

year	2001	2002	2003	2004	2005	2006	2007
# of invention patents/firm	0.016	0.034	0.038	0.072	0.070	0.095	0.117
# of utility model patents/firm	0.035	0.043	0.053	0.071	0.075	0.086	0.108
# of design patents/firm	0.052	0.068	0.065	0.094	0.088	0.086	0.097

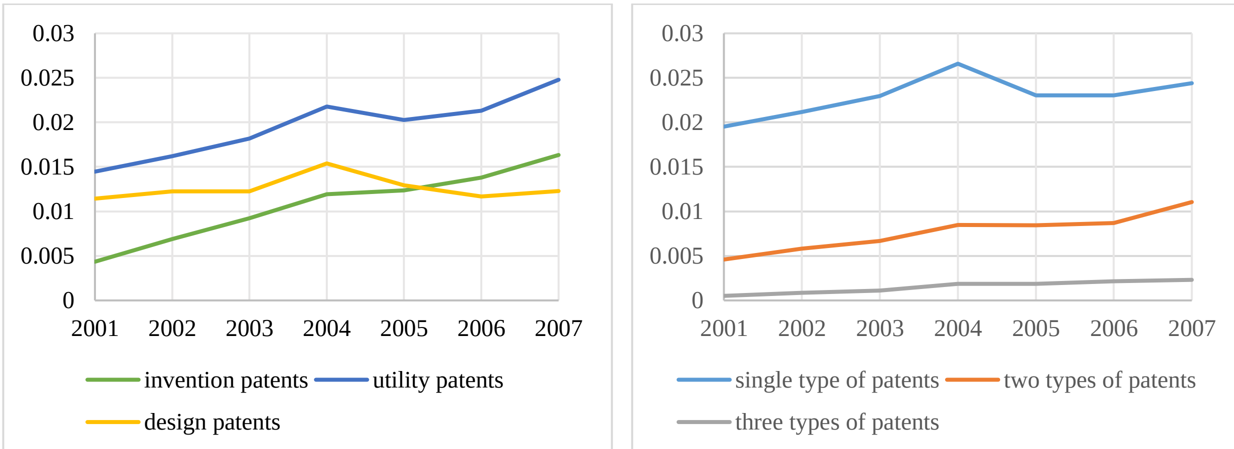
Panel C: R&D and patent application: discrete choices

		invention		utility model		design		all types	
		No	Yes	No	Yes	No	Yes	No	Yes
present R&D	No	853,150	4,249	848,283	9,116	850,982	6,417	841,760	15,639
	Yes	130,449	6,798	126,612	10,635	131,232	6,015	120,631	16,616
lagged R&D	No	606,192	3,853	602,412	7,633	604,885	5,160	597,048	12,997
	Yes	99,390	5,652	96,477	8,565	100,312	4,730	91,797	13,245

Note: matching efficiency refers the ratio of number of invention patents in the merged dataset to the published figure in China Statistical Yearbook on Science and Technology 2001-2007.

Figure 1 Trends of patents applications: extensive margin and number of patent types

Panel A: extensive margin *Panel B:* number of patent types



Notes: the vertical axis is the ratio of firms that produces patents (for a certain type) to the total number of patents.

4. Model specifications and estimation results

Models without firm-specific effects: the Poisson model and the negative binomial model

Table 2 Regression results of negative binomial mean-dispersion model with common α

Dependent Variables	**Total patents**	**Invention**	**Utility model**	**Design**
	(1)	**(2)**	**(3)**	**(4)**
Independent Variables				
$log(RD_{it})$	0.103***	0.137***	0.119***	0.074***
	(0.005)	(0.006)	(0.004)	(0.012)
$log(RD_{it-1})$	0.077***	0.089***	0.084***	0.064***
	(0.007)	(0.007)	(0.004)	(0.016)
$psub_{it}$	0.455***	0.210**	0.334***	0.584***
	(0.053)	(0.072)	(0.060)	(0.094)
FDI_{jt}	3.255***	2.257***	2.903***	4.109***
	(0.169)	(0.341)	(0.125)	(0.300)
$size_{it}$	0.706***	0.560***	0.627***	0.826***
	(0.018)	(0.026)	(0.017)	(0.035)
age_{it}	0.004*	0.002	0.002	0.005
	(0.002)	(0.003)	(0.001)	(0.003)
$htech_{it}$	0.415***	1.077***	0.289***	0.217**
	(0.041)	(0.077)	(0.042)	(0.071)
$private_{it}$	0.642***	0.467***	0.409***	0.887***
	(0.069)	(0.063)	(0.074)	(0.116)
$foreign_{it}$	0.969***	0.675***	0.553***	1.310***
	(0.082)	(0.094)	(0.079)	(0.138)
$contant$	-6.984***	-7.232***	-7.116***	-9.008***
	(0.149)	(0.175)	(0.161)	(0.271)
$time\ FE$	yes	yes	yes	yes
$ln(\alpha)$	3.653***	3.686***	3.471***	4.853***
	(0.019)	(0.051)	(0.023)	(0.028)
# of obs.	706388			
$log-pl$	-158500.4	-57374.6	-94996.8	-71006.1
χ^2	8026.2	6269.5	10592.0	2314.4
$\hat{\alpha}$	38.59	39.89	32.16	128.1

Note: the dispersion parameter α is common to all observations. χ^2 is the chi-square testing statistic under the null hypothesis that a constant-only model does better. $log-pl$ is the log-pseudo likelihood. All standard errors are robust to some kinds of misspecification recommended by Cameron and Trivedi (2009). $\hat{\alpha}$ is parameter capturing the over-dispersion.

*** 0.1% significance level; ** 1% significance level; * 5% significance level.

Table 3 Estimation results of negative binomial mean-dispersion model with α parameterized

Dependent Variables	**Total patents**	**Invention**	**Utility model**	**Design**
	(1)	**(2)**	**(3)**	**(4)**
Independent Variables				
$log(RD_{it})$	0.098***	0.132***	0.112***	0.074***
	(0.005)	(0.007)	(0.004)	(0.008)
$log(RD_{it-1})$	0.068***	0.083***	0.080***	0.053***
	(0.005)	(0.007)	(0.004)	(0.009)
$psub_{jt}$	0.376***	0.168*	0.337***	0.417***
	(0.053)	(0.069)	(0.059)	(0.095)
FDI_{jt}	3.164***	2.495***	3.040***	3.264***
	(0.160)	(0.327)	(0.138)	(0.265)
$size_{it}$	0.718***	0.620***	0.656***	0.786***
	(0.018)	(0.023)	(0.024)	(0.028)
age_{it}	0.001	-0.001	0.001	-0.001
	(0.002)	(0.002)	(0.001)	(0.002)
$htech_{it}$	0.411***	1.141***	0.297***	0.160*
	(0.038)	(0.066)	(0.045)	(0.062)
$private_{it}$	0.690***	0.465***	0.423***	0.945***
	(0.055)	(0.062)	(0.066)	(0.095)
$foreign_{it}$	0.921***	0.627***	0.527***	1.226***
	(0.069)	(0.097)	(0.076)	(0.114)
$time\ FE$	yes	yes	yes	yes
Dependent variable: ln α				
$size_{it}$	-0.392***	-0.380***	-0.328***	-0.509***
	(0.011)	(0.029)	(0.018)	(0.015)
age_{it}	-0.011***	-0.006**	-0.014***	-0.002
	(0.001)	(0.002)	(0.002)	(0.001)
$htech_{it}$	-0.879***	-0.616***	-0.315***	-1.016***
	(0.033)	(0.075)	(0.049)	(0.047)
$private_{it}$	-0.253***	-0.359***	-0.281**	-0.420***
	(0.052)	(0.086)	(0.091)	(0.074)
$foreign_{it}$	0.130*	0.798***	0.003	-0.222**
	(0.056)	(0.099)	(0.097)	(0.080)
$time\ FE$	yes	yes	yes	yes
# of obs.	706388			
$log-pl$	-155354.3	-56215.7	-93913.4	-69304.0
χ^2	9571.3	6772.2	10421.9	2995.6

Note: α is parameterized as a function of $age, size, htech$, and $ownership$ and $year$ dummies χ^2 is the chi-square testing statistic under the null hypothesis that a constant-only model does better. $log-pl$ is the log-pseudo likelihood. All standard errors are robust to some kinds of misspecification recommended by Cameron and Trivedi (2009). $\hat{\alpha}$ is parameter capturing the over-dispersion.

*** 0.1% significance level; ** 1% significance level; * 5% significance level.

Table 4 Estimation results of the zero-inflated Poisson and negative binomial models

patent types	**Invention**		**Utility model**		**Design**	
	Poisson	*NB*	*Poisson*	*NB*	*Poisson*	*NB*
$log(RD_{it})$	0.125^{***}	0.131^{***}	0.030^{**}	0.108^{***}	-0.003	0.076^{***}
	(0.027)	(0.007)	(0.011)	(0.004)	(0.007)	(0.006)
$log(RD_{it-1})$	0.017	0.084^{***}	0.019	0.079^{***}	0.003	0.049^{***}
	(0.027)	(0.007)	(0.011)	(0.004)	(0.006)	(0.006)
$psub_{jt}$	0.533^{**}	0.267	-0.008	0.189^{***}	0.363^{***}	0.814^{***}
	(0.194)	(0.139)	(0.050)	(0.056)	(0.059)	(0.076)
FDI_{jt}	4.288^{***}	3.075^{***}	0.888^{***}	0.319	0.223	-0.470
	(0.428)	(0.447)	(0.149)	(0.262)	(0.170)	(0.268)
$size_{it}$	0.805^{***}	0.268^{***}	0.490^{***}	0.389^{***}	0.452^{***}	0.357^{***}
	(0.061)	(0.046)	(0.038)	(0.034)	(0.027)	(0.028)
age_{it}	-0.019^{***}	-0.004	-0.010^{***}	-0.014^{***}	-0.009^{***}	-0.012^{***}
	(0.003)	(0.004)	(0.001)	(0.002)	(0.002)	(0.003)
$private_{it}$	0.057	0.200	0.358^{***}	0.303^{**}	0.497^{***}	0.732^{***}
	(0.263)	(0.136)	(0.079)	(0.112)	(0.093)	(0.124)
$foreign_{it}$	-0.569	1.403^{***}	0.354^{***}	0.574^{***}	0.535^{***}	0.969^{***}
	(0.387)	(0.183)	(0.082)	(0.131)	(0.099)	(0.141)
$hitech$	0.920^{***}	0.578^{***}	0.282^{**}	0.153^{*}	-0.305^{***}	-0.683^{***}
	(0.126)	(0.115)	(0.092)	(0.072)	(0.070)	(0.079)
$log(\alpha)$	--	2.581^{***}	--	2.669^{***}	--	3.476^{***}
		(0.046)		(0.053)		(0.050)
$log-pl$	-105770.9	-56429.5	-123183.4	-93710.7	-122782.6	-68700.1
χ^2-statistic	255719.7	12490.2	39774.8	10394.1	26176.1	2709.9
Vuong test	20.54	11.62	40.56	13.87	35.71	21.83
# of obs.	706388					

Note: Vuong test is the model specification test on zero-inflated negative binomial model versus standard negative binomial model (Vuong, 1989), with the null hypothesis that the standard negative binomial model fits the data better. Year dummies are include in all specifications. The standard errors are adjusted for the correlation between equations.

*** 0.1% significance level; ** 1% significance level; * 5% significance level.

5. Dealing with Endogeneity Issues

Negative binomial model with fixed effects

Table 5 Estimation results of fixed effects negative binomial model

Dependent Variables	**Total patents**	**Invention**	**Utility model**	**Design**
	(1)	**(2)**	**(3)**	**(4)**
Independent Variables				
$log(RD_{it})$	0.030***	0.019***	0.022***	0.025***
	(0.002)	(0.003)	(0.002)	(0.003)
$log(RD_{it-1})$	0.016***	0.006*	0.011***	0.015***
	(0.002)	(0.003)	(0.002)	(0.003)
$psub_{jt}$	0.068	-0.033	0.089	0.194**
	(0.038)	(0.064)	(0.052)	(0.065)
FDI_{jt}	0.386***	0.219	0.321**	0.928***
	(0.086)	(0.152)	(0.116)	(0.137)
$size_{it}$	0.139***	0.148***	0.095***	0.221***
	(0.010)	(0.019)	(0.014)	(0.017)
age_{it}	0.000	0.001	-0.000	-0.002
	(0.001)	(0.002)	(0.001)	(0.002)
$time\ FE$	yes	yes	yes	yes
# of obs.	65140	27530	42155	28087
dropped obs.:reason I	73051	73051	73051	73051
dropped obs (groups).:reason II	568197	605807	591182	605250
	(147510)	(155232)	(152223)	(155127)
$log-pl$	-55319.2	-16803.1	-30604.3	-21716.8
χ^2	1782.4	1318.7	1088.1	582.6
Hausman test χ^2	14063.63	6130.26	8296.64	3566.26

Note: Hausman test is the specification test under the null hypothesis that random-effects model and fixed-effects model have no systematic difference in coefficients. χ^2 is the chi-square testing statistic under the null hypothesis that a constant-only model does better. $log-pl$ is the log-pseudo likelihood. Reason I for dropping observations is the single observation over the sample period; reason II is the all-zero outcomes observations. All standard errors are clustered at city level. *** 0.1% significance level; ** 1% significance level; * 5% significance level.

Patent subsidy data

Table A1 Data of patent subsidies

province	starting year	province	starting year
Beijing	2000	Henan	2002
Tianjing	2002	Hubei	2007
Hebei	2002	Hunan	2004
Shanxi	2002	Guangdong	2000
Inner Mongolia	2001	Guangxi	2004
Liaoning	2002	Hainan	2001
Jiling	2004	Chongqing	2007
Heilongjiang	2001	Sichuan	2001
Shanghai	1999	Guizhou	2006
Jiangsu	2001	Yunnan	2003
Zhejiang	2001	Xizang	2004
Anhui	2003	Shanxi	2004
Fujian	2002	Gansu	--
Jiangxi	2002	Qinghan	2006
Shangdong	2006	Ningxia	2010
		Xinjiang	2002