Impact of the InnoCom Program on Corporate Innovation Performance in China: Evidence from Shanghai

Abstract
Based on the data from the Shanghai Science and Technology Enterprises survey 2011 to 2015, this paper evaluates how the InnoCom program of subsidized corporate research and development (R&D) implemented in China affects the innovation performance of beneficiary firms from both a theoretical and an empirical perspective. We first build a theoretical model to discuss firms’ responses to three preferential policies. Then, we explore the mechanism by which companies with evaluation scores exceeding a certain threshold are more likely to be certified as high and new technology enterprises that qualify for the InnoCom program, and use a fuzzy regression discontinuity design to test whether the policy increases R&D input, profit, and the number of independent intellectual property rights. After correcting for potential endogeneity problems, the result confirms a positive, significant, and lasting impact of the InnoCom program on high-tech income and the number of intellectual property rights. Meanwhile, there is no significant impact on immediate corporate innovation investment, which suggests a crowding-out effect of government direct subsidies on a company’s internal innovation investment.

JEL Classification: L5; F14; D24; O12

Keywords: regression discontinuity, InnoCom program, innovation performance

1. INTRODUCTION

Continuing innovation, diffusion, and technical improvement are widely recognized as major stimuli to national economic growth and international competitiveness in industrial, newly industrialized, and developing economies (Pavitt and Walker, 1976; Archibugi et al., 1991; Ernst and Kim, 2002; Guan and Chen, 2012). As China's industrialization has entered a mature stage, the country's leaders have focused on nurturing technology-intensive industries as
a source of future growth (Ding and Li, 2015). Corporate technological innovation has become the leading force for a country to increase its overall competitiveness. The Chinese government has introduced a series of encouragement policies to increase support for corporate research and innovation activities, but whether these policies can achieve their expected goals is debatable. As a guiding policy, how has the InnoCom program affected the innovation performance of enterprises?

Existing economic theories show that imperfect appropriability, spillovers, and uncertainty of a company's innovation output make it difficult for the company to completely internalize the benefits of R&D investment. Therefore, without external support, the equilibrium level of private resources allocated to R&D will be lower than the social optimal level (Spence, 1984). To achieve the optimal allocation of innovative resources and reduce the financing costs and information asymmetries between developers and borrowers, most countries have formulated policies or programs to support corporate R&D activities through tax reductions, fiscal subsidies, and other incentives. These policies are intended to reduce the cost of R&D expenditure for companies and thus stimulate R&D investment. Although there is an abundance of literature on policy effects, the results are mixed.

The InnoCom program in China is a policy tool that targets qualifying high-tech enterprises (HNTEs) and awards them financial support and favorable policies. As with other forms of incentive, this program seeks to overcome competitive weaknesses, such as high costs or weak business climate, promote R&D investment by offering incentives, correct for market failures in the provision of capital and risk-taking of companies, and change the image of innovation to convey a more pro-business and marketable message. According to the Enterprise Income Tax Law and the Administrative Measures for the Determination of High and New Technology Enterprises (2008) (hereinafter referred to as Administrative Measures), firms that are certified as HNTEs qualify for a reduction of corporate income tax rate to 15 percent from the original 25 percent, which is equivalent to a 40 percent tax break. Moreover, this title endows various policy benefits such as financial support. The identified high-tech enterprises are given priority in obtaining research funding and financial appropriations. The InnoCom program, financed and managed by local government, subsidizes the innovative activities of eligible HNTEs through grants (between 50,000 and 200,000 yuan) that can cover the costs of R&D projects. The
presumed mechanism behind the InnoCom program is that public incentives are expected to stimulate private research and development, and these additional induced R&D activities will be conducive to the emergence of new products and new technologies (Czarnitzki and Hussinger, 2004). However, companies have an incentive to apply for subsidies, and thus use public funds to replace private R&D investment. If a full crowding-out effect occurs, the incentive policy of high-tech enterprise certification will not lead to any improvement in corporate innovation performance. Even if there is no crowding-out effect at all, whether these additional R&D projects brought about by public incentives lead to successful results is still unclear because innovation processes are full of uncertainty caused by the constraints of technology, markets and the institutional environment. Furthermore, companies often send fake "innovation" signals to get government R&D subsidies such as relabeling other expenses as R&D expenditures (Chen et al., 2018), making the technological and economic benefits of the InnoCom program questionable.

There is abundant literature on the influence of government R&D subsidies and tax credits on corporate innovation performance based on empirical studies. For example, Girma, Görg, and Strobl (2007) find that small and medium-sized subsidies can increase R&D investments of companies, and especially that the former can generate R&D input in addition to subsidies, while excessive subsidies may squeeze out private R&D investments. Bronzini and Piselli (2016) also draw the conclusion that government subsidies have a significant impact on the number of patents, more markedly in the case of smaller firms. Cappelen et al. (2008) believe that tax credits, to a certain extent, promote the development of new products, but do not seem to contribute to the innovation or patent applications around new products. Lokshin and Mohnen (2012) study the Dutch R&D incentive program and find that the R&D costs of firms vary with tax credits which effectively promote R&D investment. Although considerable public attention has been focused on tax credits, grants and other incentives, their effectiveness, and, more generally, on the use of public funds to encourage private business, little empirical evidence has shown whether or not government programs truly benefit companies and to what extent. The influence of the innovation environment on the R&D process is related to the effectiveness of the incentive policies formed by governments (Guan and Chen, 2012).
In this paper, we evaluate the impact of the InnoCom program implemented in Shanghai on firms’ innovation performance. We choose Shanghai because it is one of the most active high-tech economic zones in China and plays an increasingly important role in national economic growth\(^1\). The results in this paper are not only of great interest to the Chinese government, but also provide useful reference for governments of other emerging economies promoting innovation development.

This paper contributes to the existing literature in several ways. First, we build a simple economic model to illustrate the effects of government support on firms’ R&D input, especially the potential crowding-out effect. Second, and most importantly, in order to avoid the endogeneity problem in the linear regression, we adopt a fuzzy regression discontinuity design with a quasi-random identification strategy to study the causal effect of the InnoCom program on enterprise innovation performance. Most studies assess whether R&D incentives have a promotional effect using standard policy evaluation methods, such as OLS, fixed effects panel regression, difference-in-differences (DID) (Chen and Gupta, 2010; Guan and Yam, 2015; Dumont, 2017; Guceri and Liu, 2015). However, innovation incentives may have a self-selection bias and reverse causality because the revealed better performance of firms receiving the promotion may be because they are better firms and better firms are more likely to get the promotion. In our case, the systematic differences between HNTEs and non-HNTEs are probably the reasons the former enjoy favorable policies, so OLS estimation will bring biased results and a fixed-effect estimation cannot fully overcome this endogeneity. In a panel setting, DID estimation relaxes the assumption of conditional exogeneity and resolves the problem of missing data by measuring outcomes and covariates for both participants and nonparticipants in pre- and post-intervention periods. The main drawback, however, rests precisely with this assumption: the notion of time-invariant selection bias is implausible for many targeted programs in developing countries. The impact of macroeconomic changes during the period on the control group and the treatment group may be different even if we can ensure the comparability of the two groups before the implementation of the policy. Under these circumstances, a simple DID may overestimate or underestimate the true effectiveness of the

---

policy, depending on the response of the treatment and control groups to common shocks in the economy. More importantly, the choice of the control group by the DID method is subjective and arbitrary. It is necessary to meet the common trend assumption and random sampling assumption when dealing with the control group and treatment group. However, the government may non-randomly select high-tech enterprises since systematic differences between HNTEs and non-HNTEs are probably the reasons the former enjoy favorable policies, so using the DID method will lead to estimation bias. Therefore, we tackle the problem of endogeneity by using a fuzzy regression discontinuity design (Lee and Lemienx, 2010) which allows us to compare HNTEs and non-HNTEs around the cut-off and use this specific difference to infer the true effectiveness of the policy.

Furthermore, we evaluate the InnoCom program from both the input side and output side of the innovation process. In particular, we analyze the effects as presumed in the InnoCom program at the firm level: the link between HNTEs certification and R&D input to identify whether the program has directly increased the R&D input intensity of the company by leveraging the government funding, on the one hand, and indirect effects of HNTE’s certification on technological performance (i.e. innovation output) on the other hand. Technological performance is measured by new and high technology products’ (services’) income and the number of independent intellectual property rights attained, because they are the most widely accepted indicators of countries’ innovation potential and are in accordance with the standards in the Administrative Measures.

The rest of the paper is organized as follows. Section one provides a description of the incentives and discusses the importance of the government’s R&D policy. The objectives and contributions of this paper are also outlined. Section two develops a theoretical model of R&D activities. Section three describes the major policies and support mechanisms of the InnoCom program. Section four introduces the data and relative variables. Section five provides the empirical strategy. Section six presents the estimation results and robustness tests. Section seven concludes the paper.
2. THEORETICAL RELATIONSHIP BETWEEN SUBSIDIES AND R&D INPUTS

This section develops and discusses a very simple and stylized model of innovation investment whereby firms respond to three preferential policies. In this model, firms engage in an optimal level of research input (which includes both unskilled labor but also human capital input). The amount of research input is related to the firm’s costs, which in turn increases the success probability of innovation. This model yields the following intuitive result: a decrease in the firm's cost of research increases the amount of R&D input in equilibrium and if the firm inputs more, it will have a higher probability of making higher profits. Before considering how the InnoCom program affects a firm's research investment choices, we first consider a simpler setup with no such government program.

R&D Activity

To simplify the analysis, we make the following two assumptions:

Assumption 1: There is no material capital investment in R&D activities (Romer, 1990). R&D activities require not only unskilled labor but also human capital input (skilled labor). Although the two cannot be separated completely, we separate them in order to clarify the different roles of the two types of labor input in R&D activities.

Assumption 2: There is no memory in the technology: the innovation rate depends only upon the current flow of input to research, not upon past research.

In addition, we do not distinguish different forms of innovation and types of product, which avoids complex discussions on the various forms of innovation and different trading behaviors, since this is not the main focus of our analysis.

R&D activity produces a random sequence of innovations. The success rate of innovations \( I \) in the economy at any instant is \( \phi(I) \) in equation (1):

\[
\phi(I) = \phi(ln(AL^\alpha H^{1-\alpha})) = \phi(lnA + \alpha lnL + (1 - \alpha)lnH), 0 < \alpha < 0.5
\]

where \( \phi \) is a cdf function, so it is between zero and one and increasing with \( I, L \) and \( H \) are the flows of unskilled and skilled labor used in research, \( A \) and \( \alpha \) are positive constant parameters, and the function of innovation is in the form of Cobb-Douglas (An, 2009).

2 There are no contemporaneous spillovers in R&D activities, that the probability is independent of the inputs of other firms. Note that \( \phi'(I) > 0, \phi''(I) \leq 0. \)
Production

The good is produced using the fixed quantity K of physical capital, and the production function can be written as:

\[ Y = F(L, H, K). \] \hspace{1cm} (2)

We consider two types of firm: innovation leaders and followers. For convenience, we set a unit marginal cost of the innovation leader through time (Davidson and Segerstrom, 2000), and we set the marginal cost of followers as \( c, c > 1 \). When technology diffusion has not reached the level that other companies can imitate, innovating companies have absolute pricing power in the market, and followers cannot obtain excess profits. Therefore, the leader can set the price at the level of \( c \), which is the lowest price that the followers can accept. Thus, the leader earns the profit flow:

\[ R = \left( \frac{c - 1}{c} \right) \theta C, \hspace{0.2cm} 0 < \theta < 1, \] \hspace{1cm} (3)

where \( C \) is the total spending of consumers and \( \theta \) is the market share of the innovation leader. The reason why the leader does not set prices higher than \( c \), is that the leader company cannot completely occupy the entire market with the constraints of expansion barriers of the company's products, consumer cognitive delay and other factors. Instead of spending a lot of fixed costs to occupy unfamiliar new markets, companies prefer patent fees, similar to the market's franchises and chain operation models. Besides, the patents of the leader company have a period of validity, and that is why followers are willing to operate without profits at this stage.

The leader's objective is to maximize the expected present value of profits over the current interval. The firm's value function takes the form:

\[ V = \int_0^{\tilde{t}} e^{-rt} \phi(I)R dt = \frac{\phi(I)R}{r} \left( 1 - e^{-r\tilde{t}} \right). \] \hspace{1cm} (4)

where \( r \) is the discount rate and \( \tilde{t} \) is the length of time that the company controls the market after the success of innovation. Equation (4) identifies \( V \) as the future flow of output discounted at the fixed rate \( r \).

Firms' Maximization Problem

Assume there are no contemporaneous spillovers in the research process (Aghion and Howitt, 1992), that is, a firm employing the amounts \( L \) and \( H \) of the two factors in R&D will
achieve innovations with a probability of $\phi(I)$, independently of the input of other companies. By (4) and Assumption 1, the objective of a firm in choosing $L$ and $H$ is to maximize the flow of expected profits from research:

$$\max [V - (w_L L + w_H H)],$$

(5)

where $w_L$ and $w_H$ are the wages of common labor and specialized human capital, respectively, and $w_L < w_H$. The first-order condition of profit maximization with respect to $L$ and $H$, together with Equation (1) produces:

$$\frac{\partial V}{\partial L} = \frac{(1 - e^{-rt}) R \phi'(I) \alpha}{r L} = w_L \text{ and } \frac{\partial V}{\partial H} = \frac{(1 - e^{-rt}) R \phi'(I)(1 - \alpha)}{r H} = w_H.$$  

(6)

From this, we can calculate the optimal level of innovation input without any favorable policy:

$$L^* = \frac{(1 - e^{-rt}) R \phi'(I) \alpha}{rw_L} \text{ and } H^* = \frac{(1 - e^{-rt}) R \phi'(I)(1 - \alpha)}{rw_H}.$$  

(7)

The model reflects the fact that in practice the primary uncertainty of research activities from the firm’s perspective is the risk that the activity will fail. We use the literature on innovation arrival rate and incentive forms to argue that this risk is also affected by the size of the input. While our analysis is related to this prior literature, we exploit the rich differences in the forms of subsidy particular to high-tech enterprises. Before we formulate our hypotheses about the effect of policies on innovation performance, we note a few important details on the preferential treatments made by national and provincial governments for HNTEs. Depending on the forms of subsidy, preferential terms can be classified into three categories: R&D cost deductions, R&D funding, and tax credits. The first two vary with the location of the company, and the last one is a national discount policy. R&D funding is provided to motivate companies to get involved in the program and enhance the technological prowess of the economy (Czarnitzki and Fier, 2002). Policy makers intend to strengthen innovation activities by these three policies, while tax credits are regarded as the most important instrument for innovation policy because it allows for flexible responses to new challenges. We analyze these three types of subsidy separately.
**R&D Cost Deductions**

Local administrators give different proportions of R&D cost reductions to the qualified companies, and then the problem of maximizing the profits of R&D activities for innovative companies is:

\[
max[V - (1 - s)(w_L L + w_H H)] ,
\]

where \( s \) is the ratio of R&D expense subsidies given by policy makers to companies, and \( 0 < s < 1 \).

The enterprise maximizes profits through choosing optimal labor and human capital investment, so we derive:

\[
L^*_{RCD} = \frac{(1 - e^{-rt})R\phi(I)\alpha}{r(1-s)w_L} \quad \text{and} \quad H^*_{RCD} = \frac{(1 - e^{-rt})R\phi(I)(1-\alpha)}{r(1-s)w_H} .
\]  

(9)

Since \( 0 < s < 1 \), we have \( L^*_{RCD} > L^* \) and \( H^*_{RCD} > H^* \). R&D cost reductions encourage firms to input more into R&D. According to Equation (9), an increase in the degree of R&D cost deductions increases the stationary equilibrium amount of labor input \( L \) and human capital input \( H \). According to Equations (1) and (4), given fixed values of the parameters \( A \) and \( \alpha \), the increase of labor and human capital will also bring higher innovation success probability and hence higher enterprise value.

**R&D Funding**

Another favorable policy is government R&D funding. Assume the firm receives a subsidy of \( S \) for its R&D projects, in which \( a \cdot S \) is used to increase the labor input, and \( (1 - a) \cdot S \) is used to increase human capital investment. The new investments in labor and human capital of the enterprise are:

\[
\bar{L} = L + \frac{aS}{w_L} \quad \text{and} \quad \bar{H} = H + \frac{(1 - a)S}{w_H} .
\]  

(10)

The new rate of innovation becomes:

\[
\phi(\bar{I}) = \phi(ln(AL^\alpha \bar{H}^{1-\alpha})) = \phi(lnA + \alpha lnL + (1-\alpha)lnH).
\]

(11)

The firm chooses \( L \) and \( H \) to maximize its profit:

\[
max[\bar{V} - (w_L \bar{L} + w_H \bar{H}) + S] .
\]  

(12)

Using similar reasoning, the corresponding first-order condition of maximizing expected profits in R&D races are:
\[
\frac{\partial \tilde{V}}{\partial L} = \frac{(1 - e^{-rt})\phi'(I)R\alpha}{r \left( L + \frac{Sa}{w_L} \right)} = w_L \text{ and } \frac{\partial \tilde{V}}{\partial H} = \frac{(1 - e^{-rt})\phi'(I)R(1 - \alpha)}{r \left( H + \frac{S(1 - a)}{w_H} \right)} = w_H. \tag{13}
\]

The enterprise internal inputs are given by
\[
L = \frac{(1 - e^{-rt})\phi'(I)R\alpha - rSa}{rw_L} \text{ and } H = \frac{(1 - e^{-rt})\phi'(I)R(1 - \alpha) - rS(1 - a)}{rw_H}. \tag{14}
\]

Putting the last two equations into Equation (10), the number of total investments after receiving subsidies will be
\[
\tilde{L} = L + \frac{Sa}{w_L} = \frac{(1 - e^{-rt})\phi'(I)R\alpha - rSa + rSa}{rw_L} = \frac{(1 - e^{-rt})\phi'(I)R\alpha}{rw_L}, \tag{15}
\]
\[
\tilde{H} = H + \frac{S(1 - a)}{w_H} = \frac{(1 - e^{-rt})\phi'(I)R(1 - \alpha) - rS(1 - a) + rS(1 - a)}{rw_L} = \frac{(1 - e^{-rt})\phi'(I)(1 - \alpha)}{rw_L}. \tag{16}
\]

Higher innovative R&D subsidy S leads to lower enterprise internal research input, whereas the number of total investments stays the same. It is worth noting that direct government subsidies generate a complete crowding-out effect, that is, firms will use public funds to replace private R&D investment.

**Tax Credits**

Drawback schemes allow qualified firms to enjoy a lower average tax rate. We assume the firm faces a tax credit \(\Delta\). This implies that we can write the expected value as follows:
\[
V = \int_0^{\tilde{t}} e^{-rt} \phi(I)Rdt = \frac{\phi(I)R}{r(1 - \Delta)} \left( 1 - e^{-rt} \right). \tag{17}
\]

The firm aims to maximize profit by choosing \(L\) and \(H\). The optimization problem is
\[
\max [V - (w_L L + w_H H)]. \tag{18}
\]

The optimal level of research input \(L^*\) and \(H^*\) satisfies the conditions
\[
\frac{\partial V}{\partial L} = \frac{(1 - e^{-rt})\phi'(I)R\alpha}{rL(1 - \Delta)} = w_L \text{ and } \frac{\partial V}{\partial H} = \frac{(1 - e^{-rt})\phi'(I)R(1 - \alpha)}{rH(1 - \Delta)} = w_H. \tag{19}
\]

The inputs can be expressed as
\[
L^*_{\text{tax}} = \frac{(1 - e^{-rt})\phi'(I)R\alpha}{r(1 - \Delta)w_L} \text{ and } H^*_{\text{tax}} = \frac{(1 - e^{-rt})\phi'(I)R(1 - \alpha)}{r(1 - \Delta)w_H}. \tag{20}
\]
Since $0 < s < 1$, we have $L^*_{\text{tax}} > L^*$ and $L^*_{\text{tax}} > H^*$. Note that for both types of R&D investment, the profit-maximizing input of R&D effort is an increasing function of tax credits, and the increase of labor and human capital also increases innovation success probability and enterprise value.

In short, for the proportional R&D cost deductions, the government will provide a certain percentage of R&D expenses to cover the costs of projects. Unlike direct subsidies, companies need to invest a certain amount of money before they receive the corresponding proportion of subsidy. These features make policies more likely to generate positive incentive effects; the more capital that companies invest, the more subsidies they receive. The incentive mechanism of tax credits is similar to proportional R&D cost deductions, as both require companies to invest in advance to get preferential treatment.

In contrast to the discussion above, there are sizable negative incentive effects associated with public funding. Scholars have obtained evidence of crowding-out effects through theoretical models and empirical studies. For example, Wallsten (2000) finds that the government funding of firms’ R&D will totally squeeze out business R&D expenditure by using a U.S. sample. David et al. (2000) review literature about government-funded R&D and find that about one-third of the literature supports the crowding-out effect of government public funding of R&D.

One form of crowding-out is that projects with higher success probabilities and high private rates of return could have been financed by the firm either from internal or external funds, suggesting that the research subsidies are in fact superfluous and may be crowding-out private R&D resources (Lach, 2002). The subsidy is “crowding out” an investment that would otherwise be firm expenditure because government subsidies reduce R&D risks and capital costs (Lee and Cin, 2010). Companies can transfer some of their own funds from projects that are profitable but risky to productive areas after receiving subsidies, so that they can use less funds to obtain project benefits and avoid risks, which in turn produces the crowding-out effect.

In summary, we argue that the principle and means of the preferential policy affect the optimal level of research activity input, which in turn affects the flow of profits. Specifically, based on our simple model and the discussions above, we make the following two theoretical claims: First, an increase in the R&D cost deductions and tax credits increases the stationary
equilibrium amount of research input, which will also bring about higher probability of innovation and increase the enterprise value. Second, direct government funding generates a crowding-out effect, that is, firms will use public funds to replace private R&D investment. Note that the three preferential policies have different effects on innovation input and output. Therefore, the ultimate effect of the determination of HNTEs is still uncertain and needs further verification in the empirical part.

3. BACKGROUND OF THE INNOCOM PROGRAM

Different from incentive plans in many developed countries, China implements the InnoCom program which targets high-tech enterprises. This project consists of not just simple fiscal subsidies and tax credits, but also an incentive policy that encourages enterprises to innovate and improve their core innovation capabilities. According to the Authentication Measures and Guidelines on the Administration of Determination of High and New Tech Enterprises (2008), the most important determinants of an HNTE are detailed in Table 1.

Table 1. Requirements for HTNE certification

<table>
<thead>
<tr>
<th>A weighted score of four indicators</th>
<th>Number of independent intellectual property rights(^3) (30 points)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Capacity of transformation of scientific and technological achievements (30 points)</td>
</tr>
<tr>
<td></td>
<td>Level of organization and management of research and development (20 points)</td>
</tr>
<tr>
<td></td>
<td>Growth in sales and total assets (10 points +10 points)</td>
</tr>
<tr>
<td>R&amp;D Intensity</td>
<td>If the sales revenue &lt; 50 million yuan ≥6%</td>
</tr>
<tr>
<td></td>
<td>If the sales revenue is 50-200 million yuan ≥4%</td>
</tr>
<tr>
<td></td>
<td>If the sales revenue &gt; 200 million yuan ≥3%</td>
</tr>
<tr>
<td>Income</td>
<td>Revenue from high-tech products (services) of its total revenue ≥60%</td>
</tr>
<tr>
<td>Human capital input</td>
<td>Percentage of staff with advanced degrees ≥30%</td>
</tr>
<tr>
<td></td>
<td>Percentage of R&amp;D staff ≥10%</td>
</tr>
</tbody>
</table>

To qualify for the program, firms have to apply and submit to a special audit and satisfy all four conditions in Table 1. By the end of 2017, the number of certified high-tech enterprises in Shanghai had exceeded 7,000. However, there may be problems with certification. A certain amount of unqualified companies obtain certification in a variety of ways to get the policy preferences, so the implementation of this policy has been questioned by many parties. People's

\(^3\) The number of independent intellectual property rights is the sum of the number of patents, software copyrights and layout designs of integrated circuits a firm owns.
Daily Online, China News, Southern Metropolis Daily, and Economic Information Daily have published articles on the phenomenon of “pseudo-high-tech” many times, and there are a plethora of high-tech enterprise certification agencies. Nevertheless, research on the impact of high-tech enterprise certification on the innovation ability of Chinese enterprises is still largely absent. Whether this policy has promoted the industrial development of high-tech enterprises in China and promoted the core innovation of Chinese enterprises is still in question. This paper contributes to this stream of research by using microdata to examine the impact of high-tech enterprise certification on the innovation performance of companies.

The mechanism of the InnoCom program allows us to compare the innovation performance of HNTEs and non-HNTEs close to the threshold score, using a fuzzy regression discontinuity design (Hahn et al., 2001; Lee and Lemieux, 2010). Firms that score above a certain level in an assessment by a technical committee would be certified, so we can estimate the policy effect using the quasi-randomness of the assignment of HLTE certifications around the cut-off, which makes HNTEs and non-HNTEs around the threshold comparable.

4. DATA

The analysis is based on the dataset provided by Shanghai Science and Technology Committee, which gives us information about medium-and small-size science and technology enterprises, including company code, number of independent intellectual property rights, R&D investment, number of employees, and whether the company is an HNTE.

Although the statistical information includes whether the enterprise has been certified as a HNTE or not, there is no specific date as to when the company was identified and the number of new and high technology transformation achievements is not reported, which is also an important assessment indicator. To make up for this deficiency, we use the relevant information published by the Shanghai Science and Technology Committee (STCSM) 6, which provides an HNTE list every year and information on Science and Technology achievement transformation. Due to the discontinuity of the data, we pool together matching data from between 2011 and

---

4 According to the article “Pseudo-high-tech companies gathered in the Growth Enterprises Market, 27 companies enjoy tax benefits of 261 million yuan”, from January 1 to March 21, 27 companies of 148 HTEs in GEM did not meet the conditions for high-tech enterprise certification. (Economic Information Daily March 30, 2012)

5 Both HTEs and non-HTEs in Shanghai.

6 http://www.stcsm.gov.cn/
2015, and study the performance of enterprises that were identified as HNTEs in 2012. We calculate grades according to the four evaluation criteria in the *Guidelines on the Administration of Determination of High and New Tech Enterprises (2008)*, as listed in the first line of Table 1. Enterprises scoring more than 70 points meet the HNTE certification standards\(^7\). As introduced in the last section, all four conditions (a weighted score of four indicators, R&D Intensity, income and human capital input) must be met simultaneously, otherwise the score will be zero.

Table 2. Enterprise descriptive statistic reports descriptive statistics of samples with scores (original score-70) above zero and around the cutoff point. Given that our empirical strategy is based on the score assigned to each company, we have to exclude those companies that have not received a score because their enterprises did not satisfy at least one of the requirements. Note that the strategy is based on the discontinuity test around the cutoff point, and omitted companies would have received an overall score far away from the threshold, so we think their exclusion does not bias our results (Bronzini and Piselli, 2016). The remaining number of companies is 1,725, of which 764 are HNTEs. We report input and output variables that we used to construct measures of enterprise performance, including enterprise economic performance, independent intellectual property rights, material capital input, human capital input and some other variables. On average, the income of new and high technology products (services) exceeds 90 million yuan, the number of independent intellectual property rights is 4.47, the number of staff with a bachelor’s degree or above is 37.65, and the number of researchers is 32.77.

The regression discontinuity estimation relies on the assumption that the treatment is random near the cutoff, so that firm characteristics should not differ significantly just below and above the threshold. The penultimate and antepenultimate columns report the summary statistics for 5 points around the cutoff point. We also apply the t statistic to test the hypothesis that their means are equal, above and below the cutoff, and provide p-values in the last column. There are no significant differences in the number of enterprises’ patents, invention patents, software copyright registrations, layout designs of integrated circuits, enterprise internal R&D expenditure, total employees, total assets and total liabilities between the left and right side of the cutoff since the p-values of the t test are all above 0.5, meaning we cannot reject the null

\(^7\) However, we don't have the data about level of organization and management of research and development. Judging from the audit mechanism, we believe that scores are closely related to the application materials, so we set the same scores (20 points) to all enterprises.
hypothesis. By contrast, the difference is more pronounced in terms of the company's new and high technology products (services) income. It rises more than threefold as the score increases, which implies that the certification may have a certain impact on business performance. Overall, these findings support our empirical strategy.
Table 2. Enterprise descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Indicator</th>
<th>Symbol</th>
<th>All (score&gt;0)</th>
<th>Cutoff-5</th>
<th>Cutoff+5</th>
<th>t test (H0: diff = 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise economic performance</td>
<td>Total income</td>
<td>Inc</td>
<td>68,522.14</td>
<td>106,125.10</td>
<td>60,823.28</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>New and high technology products (services) income</td>
<td>HTInc</td>
<td>9,310.91</td>
<td>2,976.44</td>
<td>9,442.55</td>
<td>0.10</td>
</tr>
<tr>
<td>Independent intellectual property rights</td>
<td>Pat (Invention patents)</td>
<td>Pat</td>
<td>2.28</td>
<td>2.28</td>
<td>2.8</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>SC (Software copyright registrations)</td>
<td>SC</td>
<td>1.50</td>
<td>2.61</td>
<td>2.875</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>LD (Layout designs of integrated circuits)</td>
<td>LD</td>
<td>0.04</td>
<td>0.14</td>
<td>0.075</td>
<td>0.63</td>
</tr>
<tr>
<td>Material capital input</td>
<td>Government R&amp;D funding</td>
<td>GRD</td>
<td>902.63</td>
<td>2,258.74</td>
<td>9,989.15</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>Enterprise R&amp;D expenditure</td>
<td>ERD</td>
<td>5,988.33</td>
<td>8,603.27</td>
<td>9,571.725</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>Outsourcing R&amp;D expenditure</td>
<td>ORD</td>
<td>175.58</td>
<td>192.34</td>
<td>464.025</td>
<td>0.36</td>
</tr>
<tr>
<td>Human capital input</td>
<td>Doctors</td>
<td>Doc</td>
<td>1.02</td>
<td>1.20</td>
<td>2.38</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>Masters</td>
<td>Mas</td>
<td>7.80</td>
<td>10.19</td>
<td>17.725</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Undergraduates</td>
<td>Und</td>
<td>28.83</td>
<td>33.45</td>
<td>39.875</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>R&amp;D staff</td>
<td>RDS</td>
<td>32.77</td>
<td>38.70</td>
<td>63.65</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Total employees</td>
<td>Emp</td>
<td>91.25</td>
<td>104.57</td>
<td>126.2</td>
<td>0.58</td>
</tr>
<tr>
<td>Others</td>
<td>Total Assets</td>
<td>Asset</td>
<td>95,650.30</td>
<td>126,916.40</td>
<td>171684.60</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>Total Liabilities</td>
<td>Liab</td>
<td>52,940.58</td>
<td>67,287.40</td>
<td>59943.65</td>
<td>0.85</td>
</tr>
<tr>
<td>Indicator scores</td>
<td>The score of independent intellectual property rights</td>
<td>Sip</td>
<td>8.18</td>
<td>15.04</td>
<td>14.12</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>The score of capacity of transformation</td>
<td>Scot</td>
<td>29.18</td>
<td>21.33</td>
<td>29.09</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>The score of growth in sales</td>
<td>Ss</td>
<td>8.15</td>
<td>9.32</td>
<td>9.29</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>The score of growth in total assets</td>
<td>Sta</td>
<td>8.40</td>
<td>9.48</td>
<td>9.75</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Notes: This table reports firms’ characteristics in 2012. The unit of enterprise economic performance, material capital input, and others is thousands of RMB yuan and “Independent intellectual property rights” is measured in “number of independent intellectual property rights that the firm has achieved over the past year”. N is the number of observations with at least one non-missing value for the variable listed.
5. EMPIRICAL STRATEGY

The purpose of this study is to objectively evaluate the impact of HNTE certification on firms’ performances. Lee (2008) believes that in the absence of random experiments, regression discontinuity (RD) can avoid the endogeneity problem of parameter estimation, so it can reflect the causal relationship between variables. The key assumption of the RD method is that individuals around the cutoff have similar characteristics, which can be tested by statistical analysis. And another general assumption is that the individuals must not be able to precisely control the running variable. We think that in our case such assumptions hold, because the score is obtained by evaluating the companies on the basis of the scoring criteria since the real scores are not observed in our data and it is highly unlikely that firms can perfectly control the score for every category. In Section 6 we carry out some robustness checks to test the validity of the RD design. Thus, the discontinuity of the target variable at the threshold can be attributed to the result of a policy.

RD can be divided into sharp regression discontinuity design (SRD) and fuzzy regression discontinuity design (FRD). According to the current mechanism of the HNTE certification, companies should first submit application materials and then the applications are reviewed by expert groups and a verification agency. Therefore, not all companies satisfying the conditions will become HNTES; the company needs to have the intent first. Meanwhile, in order to obtain the fiscal and taxation benefits, unqualified companies will try to obtain certification in multiple illegal ways, including relabeling innovation costs. Hence, the mechanism of InnoCom program only makes the possibility of being certified have a jump at the cutoff specified in the policy. In other words, the probability of treatment jumps by less than one at the threshold. This feature is consistent with FRD. The discontinuity can be seen in Figure 1, which captures key elements of the relation between the running variable (corporate score minus 70 points) and HNTE certification rates. Each point on the chart indicates the probability that companies in each score can be certified as a high-tech enterprise. We can see a clear breakpoint in the certification rate of high-tech companies around 70 points and an increasing probability of certification as we move right from the cutoff (Figure 1). Therefore, we can use the discontinuity of the rating
system to identify the causality between high-tech enterprise certification and corporate innovation performance.

Notes: The x-axis variable is (actual score-70) and “0” is the cutoff. The y-axis variable is the probability of an enterprise being certified at each specified score.

Figure 1. Running Variables and HNTE Certification Rates

The fuzzy RD design allows for a small jump from 0 to 1 in the probability of assignment to the treatment at the cutoff and is usually assumed to be:

$$\lim_{\epsilon \downarrow 0} \Pr(D = 1|X = c + \epsilon) \neq \lim_{\epsilon \uparrow 0} \Pr(D = 1|X = c + \epsilon),$$

where the dummy variable $D_i(x)$ denotes the treatment of an individual $i$ within a small neighborhood near the cutoff $c$, so that we have $D = 1$ if $X \geq c$, and $D = 0$ if $X < c$.

In the FRD model, scholars believe that the average causal effect of the experiment should be the ratio of the two differences: the difference of the dependent variable $Y$ on the regression of the covariates $X$ divided by the difference of the treatment variable $D$ at the threshold. This is:

$$\tau_{FRD} = \lim_{\epsilon \downarrow 0} \frac{\E(Y|X = c + \epsilon) - \lim_{\epsilon \uparrow 0} \E(Y|X = c + \epsilon)}{\lim_{\epsilon \downarrow 0} \E(D|X = c + \epsilon) - \lim_{\epsilon \uparrow 0} \E(D|X = c + \epsilon)}.$$  \hspace{1cm} (22)

Hahn et al. (2001) point out that when the treatment effect changes in units, instrumental variables can be used to explain the FRD design, which is consistent with the view of Imbens.
and Angrist (1994). This strategy relies on a general assumption: individuals do not have precise control over \( X \), and it is necessary to assume that whether \( X \) crosses the cutoff \( c \) (the instrument) has no impact on the outcome variable except by influencing \( D \) (Lee, 2008). We think that in our case this assumption holds because it is impossible for firms participating in the InnoCom program to perfectly control the score, which is a comprehensive evaluation of four categories.

The FRD leads naturally to a simple two-stage least square (2SLS) estimation strategy (Imbens and Lemieux, 2008, Lee and Lemieux, 2010), so we can use the exogenous assignment mechanism to identify the impact of certification on a firm's innovation performance. Since the treatment depends on whether the running variable exceeds the cutoff point, we can use whether the company scores over 70 points as an instrumental variable and limit the sample to a small neighborhood around the threshold. In order to test for the discontinuity at the cutoff point, parametric and non-parametric methods have been proposed. We estimate the effect of the policy using the following parametric polynomial discontinuity regression model:

\[
T_i = \alpha_0 + \alpha_1 D_i (score_i - c > 0, D_i = 1) + \alpha_2 S_i + \alpha_3 S_i^2 + \nu_i,
\]

(23)

where \( D_i \) is the excluded instrument that provides identifying power with a first-stage effect given by \( \alpha_1 \), and \( T_i \) is a dummy variable for certification. \( T_i \) is equal to 1 if enterprise \( i \) is identified as a HNTE and to 0 otherwise. Specifically, we exploit the fact that an HNTE certification is determined by \( D_i = 1(S_i \geq 0) \) where \( S_i \) is the difference between the enterprise score and the cutoff point. The second stage regression equation in this case is

\[
Y_i = \beta_0 + \beta_1 T_{it} + \beta_2 S_i + \beta_3 S_i^2 + \epsilon_i,
\]

(24)

where \( Y_i \) is the outcome variable and \( \epsilon_i \) is the error term. In order to avoid deviations caused by different innovation measures, we use three dependent variables - innovation input, innovation profit and independent intellectual property rights - which characterize firms' innovation performance. The first dependent variable indicates input of innovation, whereas the last two indicate economic and technological output of innovation. Consistent with the Guidelines on the Administration of Determination of High and New Tech Enterprises (2008), we define innovation input \((RD) = \text{Enterprise internal R&D expenditure} + 0.8 \times \text{Outsourcing R&D expenditure}\). The \( \ln RD \) is defined as the logarithm of the company's annual innovation input. Then, we measure corporate innovation output using economic performance of innovation \( (\ln HTInc) \) and number of Independent Intellectual Property rights \((IIP)\), where \( \ln HTInc \) is
defined as the logarithm of the company's annual new and high-tech product (service) revenue, and \( IIH \) is the sum of the number of patents, software copyright registrations, and layout design registrations of integrated circuits a firm owns. Multivariate innovative indicators are used to analyze the impact of the InnoCom program on Chinese firms’ innovation performance. We add polynomial terms of \( S_i \) to construct the nonlinear relationship, and report the best specification chosen by the order of polynomial that provides the minimum Akaike information criterion (AIC).

Since high-quality companies are more likely to get higher scores and to be certified, selection bias will happen if we apply OLS to estimate Equation (24) using the original value of \( T_i \). The reason why the experimental variable \( D_i \) can avoid the bias is that the variable \( D_i \) is not disturbed by the company's quality and is highly correlated with \( T_i \). In fact, whether or not a HNTE scores 70 according to the scoring mechanism is not a major concern. Our goal is to estimate companies that are truly certified as HNTEs and whether their innovation performance is affected. In order to obtain unbiased estimation of policy effect, we can use \( D_i \) as an instrumental variable for \( T_i \), because \( D_i \) can predict \( T_i \) while being unaffected by the selection bias. Therefore, if the RD setting is valid, the above 2SLS regression estimates will be consistent, which can avoid endogeneity problems caused by missing variables. In addition, we verify the validity of the RD setting by robustness checks.

6. RESULTS

This paper analyzes the relationship between HNTE certification and corporate innovation performance. According to the RD validity, we should control the non-linear continuity effect of the score above and below the threshold. In the IV/RD estimation below, this non-linear relationship is constructed by multiple order terms of \( S \). Table 3 lists the estimation results of how the rating system affects HNTE certification. The explanatory variables are whether or not they are recognized as new and high-tech enterprises. The coefficients in column (1) - (4) turn out to be positive and statistically significant at the 1 percent level in all the specifications, suggesting that scores exceeding the threshold indeed increase the probability of the company being certified.
Table 3. The fuzzy-RD first-stage estimates

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable: Whether the firm is certified as an HNTE</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>------------------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>D</td>
<td>0.688***</td>
<td>0.437***</td>
<td>0.196***</td>
<td>0.195***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>(score-70)</td>
<td>0.003***</td>
<td>0.012***</td>
<td>0.009***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.97e-05)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(score-70)^2</td>
<td>0.0001***</td>
<td>0.0004***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.75e-06)</td>
<td>(2.28e-05)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(score-70)^3</td>
<td></td>
<td></td>
<td></td>
<td>4.30e-06***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(3.36e-07)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.016***</td>
<td>0.217***</td>
<td>0.292***</td>
<td>0.233***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.004)</td>
<td>(0.007)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>1,725</td>
<td>1,725</td>
<td>1,725</td>
<td>1,725</td>
<td></td>
</tr>
<tr>
<td>R^2</td>
<td>0.387</td>
<td>0.463</td>
<td>0.468</td>
<td>0.473</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *p<0.10; **p<0.05; ***p<0.01. Standard errors in parentheses.

The results in Table 3 are from the first stage regression. In the following, we will focus on the main regression equation (the impact of the certification on the explanatory variables, i.e., the second stage). Lee and Lemieux (2010) believe that if the RD design is valid, the estimator will be consistent for the same parameter whether we add any combination of covariates or not.

Table 4 shows the 2SLS estimation results of the policy impact on innovation input and output in the current year considering three samples around the cut-off: the whole sample and sample windows of 20 and 10 scores. No other control variables were added except for whether they were identified as HNTES, the instrumental variable (score dummy variable), and polynomial terms of $S$.

Table 4. Baseline results: 2SLS estimates for HNTE certification

<table>
<thead>
<tr>
<th></th>
<th>lnRD</th>
<th>lnHTInc</th>
<th>IIP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[65,75]</td>
<td>[60,80]</td>
<td>[0.100]</td>
</tr>
<tr>
<td></td>
<td>[65,75]</td>
<td>[60,80]</td>
<td>[0.100]</td>
</tr>
<tr>
<td>$T$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.401</td>
<td>1.306</td>
<td>0.401</td>
</tr>
<tr>
<td></td>
<td>(0.844)</td>
<td>(2.502)</td>
<td>(0.844)</td>
</tr>
<tr>
<td>Cons.</td>
<td>7.547***</td>
<td>6.838***</td>
<td>7.547***</td>
</tr>
<tr>
<td></td>
<td>(0.167)</td>
<td>(1.501)</td>
<td>(0.167)</td>
</tr>
<tr>
<td>Obs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>129</td>
<td>1,063</td>
<td>1,725</td>
</tr>
</tbody>
</table>

Notes: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses.

From the result in columns (1) - (3), we can see there is no significant effect of certification.
on innovation input, although the coefficient is positive. Czarnitzki and Hussinger (2004) argue that government R&D subsidies have incentive effects on R&D investment (government subsidies and other preferential policies encourage companies to increase R&D investment), but also have a crowding-out effect (after being certified as high-tech enterprises, companies can obtain more government R&D subsidies which replace enterprise internal R&D expenditure).

As discussed in section two, R&D cost deductions, R&D funding and tax credits have different impacts on R&D input. In our case, the empirical results show that incentive policy has no significant impact on corporate R&D investment, indicating that companies do not significantly increase their R&D expenditure even if they obtain more funding or tax credits. Therefore, the policy is likely to have a limited effect on the promotion of independent innovation input, and may even cause a substitution effect or crowding-out effect on the independent innovation expenditure of a company.

Consistent with theoretical intuition, the income related to high-tech products will be affected by the certification. From column (6) of Table 4, we can see the estimated coefficient of income associated with high-tech products is 2.939, which is significantly positive at the 1 percent level, meaning that the income of products (services) related to high-tech is expected to nearly triple after the company obtains high-tech certification. The last two columns of Table 4 show that the program has a significant positive effect on the number of independent intellectual property rights of enterprises, with a coefficient of 4.637, meaning that the number of IIPs increases on average more than 4 times for firms receiving the preferential treatment. The coefficients are bigger with the sample window closer to the cut-off, becoming large and a little less plausible in the [65,75] sample. This is probably because the data size is much smaller when we estimate the jump using the sample very close to the cut-off. On the basis of the analysis above, we believe that the direct effects of the subsidy mechanism are not significant, but indirect effects play a very important role.
Notes: The x-axis variable is (actual score-70) and “0” is the cutoff. Y-axis variables are lnRD, lnHTInc, IIP of 2012 respectively.

Figure 2. Treatment Period

Figure 2 shows a graphical analysis of the outcome variables as a local linear function of the score (score-70). The graphs give us visual evidence of a discontinuity, which is weaker in the R&D input case. We can see that economic performance of innovation and the number of independent intellectual property rights have a significant jump around the threshold.

Table 5. Baseline results: 2SLS estimates for HNTEs certification (in the long run)

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>lnRD</th>
<th>lnHTInc</th>
<th>IIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.223</td>
<td>0.305</td>
<td>1.103</td>
</tr>
<tr>
<td></td>
<td>(1.701)</td>
<td>(1.552)</td>
<td>(1.884)</td>
</tr>
<tr>
<td>Cons.</td>
<td>7.113***</td>
<td>7.238***</td>
<td>6.928***</td>
</tr>
<tr>
<td></td>
<td>(0.747)</td>
<td>(0.853)</td>
<td>(0.974)</td>
</tr>
<tr>
<td>Obs</td>
<td>1,140</td>
<td>1,107</td>
<td>830</td>
</tr>
</tbody>
</table>

Notes: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses.

Table 5 reports the results of the policy’s impact on innovation input and output in 2013, 2014, and 2015. The first three columns of Table 5 report no significant increase in R&D investment by certified firms and the results confirm the existence of the crowding-out effect. Columns (4) and (5) report the increase in new and high technology products’ (services’) annual income. For the first two years after being certified, this increase is more than triple, which means that the impact on innovation output is sustainable, although the results are no longer significant after three years of being certified.

Robustness

In this section, we carry out several robustness checks to test the validity of our empirical design and the sensitivity of our results, including RD continuity assumption test, jumps at non-
discontinuity points test, involving covariates test and classification regression test. We also use a difference-in-difference method based on propensity score matching (PSM+DID) to further test the robustness of the estimated results.

**RD continuity assumption test**

A first concern about RD designs is the possibility of other changes at the same cutoff value of the covariate. In order to verify the validity of the RD setting, it is necessary to test the continuity of other characteristics of the company or control variables. Ideally, around the cutoff point, other control variables that do not directly affect the company's innovation performance through the score should be continuous. We estimate the following parametric polynomial discontinuity regression model:

\[ Z_i = \gamma_0 + \gamma_1 T_i + \gamma_2 S + \gamma_3 S^2 + \mu_i \]  

where \( Z_i \) is a characteristic variable of companies, such as the size of the company, asset-liability ratio, the number of employees, the proportion of staff with advanced degrees, and the liabilities of the company. We find no discontinuity for any individual characteristics examined and no significant results for these variables that should not be affected, as shown in Table 6, indicating that the application of a discontinuity regression design method is appropriate.

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>( \text{lnAsset} )</th>
<th>( \text{DARatio} )</th>
<th>( \text{Emp} )</th>
<th>( \text{HDRatio} )</th>
<th>( \text{lnLiab} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>1.588</td>
<td>0.898</td>
<td>73.55</td>
<td>-0.0260</td>
<td>-0.505</td>
</tr>
<tr>
<td></td>
<td>(0.976)</td>
<td>(2.425)</td>
<td>(103.9)</td>
<td>(0.141)</td>
<td>(7.114)</td>
</tr>
<tr>
<td>Cons.</td>
<td>8.385***</td>
<td>0.402</td>
<td>54.91</td>
<td>0.576***</td>
<td>7.823***</td>
</tr>
<tr>
<td></td>
<td>(0.491)</td>
<td>(0.948)</td>
<td>(55.10)</td>
<td>(0.0631)</td>
<td>(2.681)</td>
</tr>
<tr>
<td>Obs</td>
<td>1,634</td>
<td>1,634</td>
<td>1,725</td>
<td>1,725</td>
<td>1,519</td>
</tr>
</tbody>
</table>

*Notes:* *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses.

**RD random assumption test**

Another concern about RD designs is the possibility that agents can precisely control the running variable. In such a case, the treatment around the threshold is as if it were not randomized, and the impact of the program cannot be identified by the discontinuity of the outcome variable at the cutoff point (Hahn et al. 2001). Chen et al. (2018) use enterprise income tax records between 2008 and 2011 in China and find that firms can actually reclassify some
expenditures as R&D to meet the minimum requirement for one category artificially. To check whether similar bunching patterns exist in our data, we follow Chen et al. (2018) by providing descriptive evidence. Figure 3 plots the empirical distribution of the R&D intensity, the ratio of R&D staff and the ratio of staff with advanced degrees of Shanghai firms in 2012. The first panel of Figure 3 shows the histogram of the overall R&D intensity distribution and the second panel plots the histogram of R&D intensity for small size firms. From the figures, we can see there are no clear bunching patterns at 3%, 4%, and 6% of R&D intensity, which corresponds to the three thresholds. Similarly, the other two figures present the distribution of human capital input requirements of HTNE certification and exhibit no clear bunching pattern. We think that in our case the assumption that individuals have imprecise control over the assignment variable holds, because the score is a comprehensive measure of four categories and it is highly unlikely that firms participating in the program can perfectly control all requirements to just above the threshold.

Notes: The first two figures plot the empirical distribution of R&D intensity for all sizes of firm and only small firms that have R&D intensity between 0.5% and 10%. The other two figures report the proportions of R&D staff and staff with advanced degrees of the total number of employees, respectively. Note that the red lines are the thresholds that qualify a company to apply for InnoCom certification.

---

8 To qualify for the government program, firms are benchmarked on various attributes, one of which is current R&D expenditure: firms must spend a minimum amount on R&D to qualify, depending on their size. See more details in Table 1, page 12.
Falsification tests

Another test involves estimating jumps at points where there should be no jumps. As in the treatment effect literature (e.g., Imbens, 2004, Imbens and Lemieux, 2007), the approach used here consists of testing for a zero effect in settings where it is known that the effect should be 0. If the jump in innovation performance detected for treated enterprises is due to the InnoCom program, in the absence of treatment we should not find any discontinuity. Here we re-estimated the model for the company's R&D expenditure, economic performance of innovation and the number of independent intellectual property rights over the period before the program (year 2011). Figure 4 shows that before being certified there were no positive discontinuities of the functions around the cut-off.

Tests involving covariates

As discussed above, Lee and Lemieux (2009) think that if the no-manipulation assumption holds, the unobserved factors between the treatment group and the control group should be similar in principle, so the inclusion of baseline covariates, no matter how highly correlated they are with the outcome variables, should not affect the estimated results. We simply include the covariates directly, after choosing a suitable order of polynomial. Based on previous literature
(Zhou and Luo, 2005; Hirshleifer et al., 2011), combined with the practices of China, we consider the following factors that may affect innovation activities in technology companies as control variables: (1) Size of the company (Size), defined as the natural logarithm of the total assets of the company at the end of the year. (2) The debt ratio (Leverage), defined as total debt divided by total assets. (3) Return on assets (ROA), defined as profit divided by total assets. (4) Growth (Gsales) is defined as the annual growth rate of corporate sales revenue. (5) Subsidy (Subsidy), defined as the natural logarithm of annual R&D subsidies from the government.

Table 7. Robustness: add control variables

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>(1) lnRD</th>
<th>(2) lnHTInc</th>
<th>(3) IIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.885*</td>
<td>1.595***</td>
<td>3.965***</td>
</tr>
<tr>
<td></td>
<td>(0.493)</td>
<td>(0.293)</td>
<td>(0.273)</td>
</tr>
<tr>
<td>lnAsset</td>
<td>0.731***</td>
<td>0.126***</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.014)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Lev</td>
<td>0.015**</td>
<td>0.003</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>ROA</td>
<td>-0.034**</td>
<td>-0.004</td>
<td>-0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.003)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>lnGRD</td>
<td>0.124***</td>
<td>0.108***</td>
<td>0.108***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.031)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>RDSratio</td>
<td>4.174***</td>
<td>2.377***</td>
<td>0.927***</td>
</tr>
<tr>
<td></td>
<td>(0.206)</td>
<td>(0.094)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Cons.</td>
<td>-3.016***</td>
<td>-0.482***</td>
<td>-0.035</td>
</tr>
<tr>
<td></td>
<td>(0.430)</td>
<td>(0.099)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>Obs</td>
<td>1.634</td>
<td>1.634</td>
<td>1.634</td>
</tr>
<tr>
<td>R²</td>
<td>0.582</td>
<td>0.121</td>
<td>0.213</td>
</tr>
</tbody>
</table>

Notes: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses.

Table 7 shows the effect of the program on the innovation performance of companies after including firm size, debt ratio, return on assets, growth, and government subsidies. We find that the coefficients of the high-tech products (services) income and the number of independent intellectual property rights are slightly lower than those without control variables (Table 4), but the signs and significance level of the coefficients of innovation output do not differ in a significant way. There is some evidence of a slight difference in the R&D input, but only at the 10 percent level. We interpret these findings as further evidence of the positive impact of the
InnoCom program.

*Classification regression test*

There may be differences in the innovation performance of enterprises with different types of ownership (Wu, 2012). The response to incentive policies is also different. The relationship between enterprise ownership and innovation performance in the process of economic transition in China is of particular concern. State-owned enterprises and non-state-owned enterprises have their own advantages and disadvantages in terms of technological innovation. Most state-owned enterprises are powerful and have sufficient R&D funds. Non-state-owned enterprises have a higher awareness of market competition and more incentive to innovate to gain competitive advantage. Therefore, the policy may have different influences on innovation performance of different types of enterprise.

**Table 8. Robustness: classification regression test for high-tech income**

<table>
<thead>
<tr>
<th>Type of Company</th>
<th>Coeff.</th>
<th>s.e.</th>
<th>T value</th>
<th>P value</th>
<th>( R^2 )</th>
<th>Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>State-owned enterprise</td>
<td>2.895</td>
<td>1.178</td>
<td>2.460</td>
<td>0.016</td>
<td>0.11</td>
<td>72</td>
</tr>
<tr>
<td>Joint stock company</td>
<td>4.237</td>
<td>1.621</td>
<td>2.610</td>
<td>0.010</td>
<td>0.12</td>
<td>103</td>
</tr>
<tr>
<td>Limited liability company</td>
<td>2.493</td>
<td>0.138</td>
<td>18.090</td>
<td>0.000</td>
<td>0.07</td>
<td>9,480</td>
</tr>
<tr>
<td>Company limited by shares</td>
<td>1.342</td>
<td>0.531</td>
<td>2.530</td>
<td>0.014</td>
<td>0.14</td>
<td>63</td>
</tr>
<tr>
<td>Private limited liability company</td>
<td>2.364</td>
<td>0.345</td>
<td>18.090</td>
<td>0.000</td>
<td>0.08</td>
<td>1,103</td>
</tr>
<tr>
<td>Joint venture of mainland and HMT</td>
<td>1.660</td>
<td>0.763</td>
<td>2.180</td>
<td>0.033</td>
<td>0.06</td>
<td>73</td>
</tr>
<tr>
<td>HMT solely owned</td>
<td>2.539</td>
<td>0.633</td>
<td>4.010</td>
<td>0.000</td>
<td>0.09</td>
<td>184</td>
</tr>
<tr>
<td>Sino-foreign joint venture</td>
<td>1.227</td>
<td>0.421</td>
<td>2.920</td>
<td>0.004</td>
<td>0.08</td>
<td>180</td>
</tr>
<tr>
<td>Foreign-funded enterprise</td>
<td>0.942</td>
<td>0.535</td>
<td>1.760</td>
<td>0.079</td>
<td>0.03</td>
<td>434</td>
</tr>
</tbody>
</table>

*Notes:* HMT is an acronym for Hong Kong, Macau and Taiwan. The dependent variable in Table 8 is the logarithm of the company's annual new and high-tech product (service) revenue.

Table 8 reports the impact of the program on high-tech income of different types of company. Although the estimation coefficients of different types of company are different, they are all significantly greater than 0, indicating that the policy has increased the income related to high-tech products. And the promotion effects on joint stock companies, state-owned enterprises and HMT solely owned companies are the most significant.

*Difference-in-differences*

By 2015, a total of 6,071 companies in Shanghai were recognized as high-tech enterprises\(^9\).

---

\(^9\) Data source: [http://www.shanghai.gov.cn/nw2/nw2314/nw2318/nw26434/u21aw1109178.html](http://www.shanghai.gov.cn/nw2/nw2314/nw2318/nw26434/u21aw1109178.html)
which provides a good quasi-natural experiment to use the difference-in-difference method. Specifically, in our sample, enterprises that have been certified as HNTEs constitute a treatment group, and the rest of the enterprises that did not obtain certification naturally make up the control group. To maintain the consistency of the research, we still study the impact of the policy on enterprises that were certified as HNTEs in 2012. The time window includes each year before and after, so the time span is from 2011 to 2013. Thus, we run the following difference-in-differences estimation over the samples:

\[ Y_{it} = \beta_0 + \beta_1 HT_i \times Year_{t} + \beta_2 HT_i + \beta_3 Year_{t} + \beta_4 X_{it} + \gamma_t + \mu_i + \epsilon_{it} \]  

(26)

where \( Y_{it} \) is the outcome variable and \( X_{it} \) is a vector of multiple control variables for firm \( i \) at period \( t \). In keeping with the prior setting, we still choose Size, Leverage, ROA, Gsales and Subsidy as the control variables. \( Year_{t} \) is a dummy variable equal to 1 in the treatment period and zero otherwise. \( HT_i \) is a treatment variable equal to 1 if the enterprise is HNTE and zero otherwise. \( \gamma_t \) is the year fixed effect and \( \mu_i \) is the individual fixed effect. The coefficient of interest is \( \beta_1 \), which multiplies the two dummies and which is equal to 1 for those individuals in the treatment group in the treatment period. It measures the impact of the program on the company’s innovation performance. If innovative policy really improves innovation performance, then \( \beta_1 \) should be significantly positive.

The first three columns of Table 9 show the estimation results of DID. In column (1), unlike the results of FRD estimation, the “Diff-in-Diff” coefficient for the R&D investment (0.154) indicates a significant difference between the treatment group and the control group at the 5 percent level. The coefficient is positive, which indicates that the innovation input of firms in the treatment group is higher than that in the control group. Similarly, in column (2), the “Diff-in-Diff” coefficient for new and high technology products (services) income (2.914) indicates a positive, significant difference between the groups at the 1 percent level and the same for the number of independent intellectual property rights. Therefore, innovation performance is higher in the treatment group than in the control group.

The different results from FRD and DID estimation also support our view that there may be selection bias if we use traditional methods. Therefore, we combine PSM with DID methods to help resolve this problem, by matching units in the common support or overlap across the treatment and control samples. Specifically, the propensity score can be used to match
participant and control units in the base (pre-program) year, and the treatment impact is calculated across participants and matched control units within the common support. Hirano, Imbens, and Ridder (2003) show that a weighted least squares regression, by weighting the control observations according to their propensity score, yields a fully efficient estimator. Columns (4) - (9) show similar results compared with the FRD estimation. However, the interaction term of the R&D investment is positive, but not statistically significant. The results suggest that the effect of the policy is positive and significant on innovation output, but it is weaker, though positive, on R&D investment.
<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>DID</th>
<th>PSM+DID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>lnRD</td>
<td>0.154** (0.062)</td>
<td>2.914*** (0.160)</td>
</tr>
<tr>
<td>lnHTInc</td>
<td>0.422*** (0.027)</td>
<td>0.798*** (0.071)</td>
</tr>
<tr>
<td>IIP</td>
<td>9.80E-07 (9.59E-06)</td>
<td>-2.11E-05 (2.47E-05)</td>
</tr>
<tr>
<td>lnAsset</td>
<td>-0.006 (0.007)</td>
<td>-0.019 (0.019)</td>
</tr>
<tr>
<td>Lev</td>
<td>0.078*** (0.009)</td>
<td>0.0378 (0.024)</td>
</tr>
<tr>
<td>ROA</td>
<td>5.211*** (0.076)</td>
<td>3.036*** (0.195)</td>
</tr>
<tr>
<td>lnGRD</td>
<td>-0.722*** (0.252)</td>
<td>-6.668*** (0.650)</td>
</tr>
<tr>
<td>RDSratio</td>
<td>15.921 (15.921)</td>
<td>15.921 (15.921)</td>
</tr>
<tr>
<td>Cons.</td>
<td>0.380 (0.252)</td>
<td>0.157 (0.650)</td>
</tr>
</tbody>
</table>

Notes: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors are clustered by firms.
7. CONCLUSIONS

Although China has become used to supercharged rates of expansion, its economy still faces many challenges with a lack of enterprise innovation and an imbalance of the industrial structure. This paper evaluates the impact of the InnoCom program on innovation activities of new and high-tech firms based on survey data of Shanghai’s scientific and technological enterprises from 2011 to 2015. We exploit the mechanism of the program to apply an IV/RD parameter estimation method under a fuzzy discontinuity regression design framework to deal with the endogeneity problem in the regressions. Causal inferences from RD designs are potentially more credible than those from typical natural experiment strategies (e.g. difference-in-differences) (Lee, 2008). In addition, we look at the impact of the InnoCom program on both innovation input and innovation output.

We find that the InnoCom program has a positive effect on the innovative output of HNTEs, including high-tech product revenues and independent intellectual property rights, which is consistent with the results using the DID method based on propensity score matching. Another effect of the policy is insignificant impact on corporate innovation investment. The reason might be that Government subsidies are likely to have a crowding-out effect on a company’s internal innovation investment. The results of high-tech product revenues and independent intellectual property rights, read jointly, suggest that the increase in innovation output is mainly due to favorable treatment. We estimate the impact for another three years after the determination, and it turns out that the certification has a sustained impact on innovation output.

In addition, this paper develops a series of robustness tests. The variables that should not be affected by the determination system are continuous around the breakpoint and the assignment variable cannot be precisely manipulated. The counterfactual test shows no positive discontinuities of the functions around the cut-off over the pre-treatment period (before the program). After involving control variables, the coefficients of the high-tech products’ (services’) income and the number of independent intellectual property rights are slightly lower than those without control variables, but the signs and significance level do not differ in a
significant way. Our results are robust to the sensitivity exercise after considering the heterogeneity of the ownership of enterprises, and are also confirmed by combining PSM with DID methods, proving stability and reliability.

Of course, our study has a number of caveats. First, we only consider Shanghai as an example. This is a common disadvantage of RD that the interval validity is strong, but the external validity is questionable. We do not claim that our results hold for other regions or for China as a whole, but our empirical method can be directly applied to evaluate the policy effect. Second, on January 29, 2016, the Ministry of Finance, the Ministry of Science and Technology and State Administration of Taxation revised the Guidelines on the Administration of Determination of High and New Tech Enterprises which had been in place for eight years. It is possible that what we have found might only be partially valid for the year before the new policy was implemented. We leave these issues for future research.
REFERENCES


Yanbing Wu. (2012). Which type of ownership is most innovative in China.? The Journal of World Economy (6), 3—29.