Innovation and Geographic Concentration of Firms in China

張，怡

(Degree)
博士（経済学）
(Date of Degree)
2019-09-25
(Date of Publication)
2021-09-25
(Resource Type)
doctoral thesis
(Report Number)
甲第7575号
(URL)
https://hdl.handle.net/20.500.14094/D1007575

博士論文

Innovation and Geographic Concentration of Firms in China
（中国における企業のイノベーションと地理的集中）

令和元年 6 月 10 日

神戸大学経済学研究科
経済学専攻
指導教員　浜口伸明
張　怡

Acknowledge

During my doctoral study, I received a lot of help and encouragement from many people, without whom I might not have been able to get my doctorate. Here, I would like to express my appreciation to them.

First of all, I would like to thank my supervisor Prof. Hamaguchi. I completed my thesis under his encouragement guidance. Thanks for his advice and guidance to me over the past five years, teaching me how to do research, always affirming my research work, and encouraging me to make more attempts. Besides, he is so kind and willing to help others, rigorous in research and full of love for life, these deeply affected me. From him, I not only learned how to do research but also learned how to live a good life.

Secondly, I'd like to thank Prof. Kajitani. I especially appreciate his providing of the industrial enterprises' database and helpful suggestion for my research in the past five years, which benefited me a lot. I also would like to thank Prof. Shen for his useful comment.

Then, I'd like to thank my predecessor, Dr Fukui, for helping me find the code needed for my research, as well as the useful advice given in each presentation. Thanks to Mr Takano (Tsukuba University) for providing me with the code and code specification I needed for my research.

Finally, I would like to thank my parents and dear friends who have been supporting me in challenging myself for a long time. I want to express my gratitude to my parents for their financial and spiritual support to my study path and for being my most reliable support so that I can focus on my research. Thank my friends, especially Dr Guo, encourage me when I am frustrated, help me to find a way to solve when I meet with difficulties, cheer for me when I have progressed in my research, the completion of my PhD thesis without her encouragement to me.

Content

Chapter 1 Introduction

The distribution of economic activities is uneven. From a global perspective, industrial production concentrates in developed countries and emerging countries. From a more specific spatial distribution, it agglomerates in some regions of these countries. In the case of China, this phenomenon also can be observed. Since the Reform and Opening Up, economic activities and labourers increasingly concentrate in the eastern coastal areas, which has promoted the rapid economic development of the region.

Marshall (1890,1920) indicates the reasons that economic activities tend to cluster in some specific areas are the sharing of input, a pooling of particular labour, and knowledge spillover. Then, Krugman (1991) present the concept of transport cost to explain why economic activities spatially concentrate. Agglomeration of economic activities brings positive externalities but also brings adverse effects such as rising labour costs and congestion effects. Therefore, it is essential to implement policies to maintain the positive effects of agglomeration and to eliminate adverse effects.

In China, the economic gap between the eastern, central and western regions has been expanding and becoming a destabilizing factor in the development of the whole society. Therefore, the problems to be solved are: on the one hand, we must maintain the existing agglomeration effect in the eastern area to promote economic development, on the other hand, we must encourage industrial agglomeration and economic growth in the inland regions, and narrow the income gap between the east and the west. Especially for the development of the western area, not only the investment in infrastructure but also the elements which can cause economic activities, and population agglomeration should be paid attention, to achieve an endogenous economic growth.

Promoting innovation activities and creating new businesses are keys to solving these problems. This study aims to analyze the factors affecting the location of enterprises and the differences in regional innovation capabilities and the reasons for the differences. They will be discussed in the context of China. Our study gives some additions to the existing literature.

The determinants of entrants' location choice

Scholars and policymakers pay attention to the issue of creating new business because it is an essential factor in the sustainable development of the regional economy. In the literature, researchers try to use various of methods to explore the determinants of new firms' choice of their locations (McFadden, 2001; Friedman et al., 1992; Head and Mayer, 2004; Cheng and Stough, 2006; YANG et al., 2017) Most studies focus on the location choice of FDI when they enter a foreign market. For instance, Head et al. (1995) examine the agglomeration benefits affect Japanese manufacturing investments in the US in location choices by using a conditional logit model to explain how the external environment affects decision making. Since the Reform and Opening Up, many studies explore the determinants of FDI location choices in China.

Similar to prior research, studies on the determinants of location choice of FDI in China focus on the external conditions of firms such as the market potential, industrial agglomeration, externality of variety, labor cost and other factors (Chen, 2009; Jean et al., 2011; Cheng and Stough, 2006).

Some scholars take also a more considerable explanation based on the situation of China. For instance, Du et al. (2008) examine the impacts of an economic institution like property rights protection, contract enforcement on the location choice of US multinationals in China. Their study estimates a conditional logit model and the results show that US multinationals prefer to locate in regions that have better property rights protection and lower government intervention in business operation. Besides, Du et al. (2012) present a calculation method of cultural distance as a new proxy of distance, to examine how the cultural proximity plays a role on affecting the location choice of FDI from different countries.

With the growth of economy, Chinese private enterprises have also grown up. The focus of scholars' attention has expanded from FDI to all kinds of ownership enterprises in China. Shi and He (2018) explore the determinants of new firms set up in China by taking the metal products industry as an example. Their study shows that the extent of global linkage, the extent of the regional competition in a city have a positive effect on the new firms to start their business in the city. However, more research is needed to have better understand on enterprises located in China.

The determinants of regional innovative capability

Besides the creation of new business, the regional innovation capability is essential to sustain the competitive advantages of regions. Therefore, issues of regional innovative capacity are also hot and attract lots of scholars' attention. One of the issues is the determinants of regional innovative capability (Krugman, 1991; Maskell and Malmberg, 1999).

The difference in regional innovation capability is essential to explain agglomeration because firms are attracted to use the externality of knowledge spillover (Marshall, 1890). Therefore, scholars are interested in finding out factors that affect the innovative capability of regions from both theoretical and empirical methods. (Jacobs, 1969; Gleaser et al., 1992; Audrestch and Vivarelli, 1996).

In the empirical literature, based on a knowledge production function (Jaffe, 1989; Feldman, 1994; Acs et al., 2002), the most common determinants are R&D expenditure, the variety of knowledge in the region, human capital and market structure (Jaffe, 1989; Jacobs, 1969; OECD, 1995; Amable and Petit, 2001; Buesa et al., 2010).

China has a regional difference in innovation as well. Innovation capability in the eastern areas is obviously more advanced than the central and western areas. Studies explore the reason for the difference from many perspectives. Zhou et al. (2015) find the share of new product revenue to gross revenue has a positive influence on the regional innovation capability. Lei et al. (2019) adopt a factors' decomposition analysis to show that the capital and the number of R&D employers are the key factors that determined the regional innovation capability.

However, most of the previous literature on the determinants of regional innovation capability in China is studied at a provincial level due to the limitation of the data resource. More detailed research in a smaller spatial level is desirable.

In addition, these inland big cities like Xi'an, Chongqing also have a high-speed growth in innovation. Some scholars noticed this change and conducted research. Sigurdson and Polonka (2008) set their sight on Chongqing and analyze the industries of Chongqing to indicate that Chongqing is an innovative city in the western China. Further research on regional innovation in the western region is needed.

The localization of innovation activity

Regional innovation activity is the foundation of regional innovation capability. To better understand the function of innovation in regions, scholars pay attention to the characteristics of regional innovation activity. The previous empirical studies show that innovation activity is spatially concentrated (Carlino and Kerr, 2014; Buzard and Carlino, 2013; Carrincazeaux et al., 2001).

The study on the concentration of innovation activity is based on the calculation of the degree of concentration of economic activity. The most common indexes for measuring the spatial concentration are LQ index, locational Gini coefficient and EG index. Feldman and Audretsch (1999) estimate the concentration of innovation activity by using an EG index. However, these measurements have some disadvantage caused by the count-data aggregation. Most of all, the aggregated data blur the real spatial distribution of innovation activity in our impression. To solve this problem, Duranton and Overman (2005) first proposed a distance-based method by using micro-data.

Scholars are interested in the spatial concentration of innovation activity in China because of the rapid development of Chinese innovation. The existing literature shows that innovation activities are significantly concentrated in the advanced eastern provinces (Wei et al., 2011; Cao and Qin, 2012; Jiang, 2013). Ma and Liu (2019) narrow the scope of their study to the three mega-economic zones, including Beijing-Tianjin-Hebei, Yangtze River Delta and Pearl River Delta. However, their study still is based on a provincial level and lack of detail at the micro level.

Therefore, this study aims to try to complement the shortcomings of existing research from these three aspects to contribute to better understanding the innovation and geographic concentration of Firms in China.

In Chapter 2, we focus on the location choice of new firms of Yangtze River Delta. We analyze the determinants of their location decisions. We consider the external environment that affects firms' location choice following the previous studies for our hypotheses. The market potential of regions has an impact on the location choice of private firms' location choice but has no effect on the location choice of FDI. In China, the land supply also plays a role in affecting the location choice of each ownership firms. The congestion effect has a negative impact on the location choice of firms. In addition, we consider the characteristics of firms also affect their location choices such as the concentration of industries to which firms belong to, and the financial constraints faced by firms.

Because our study consider both characteristics of alternatives and characteristics of firms themselves, a mixed logit model is suitable for our analysis. By using a dataset of 1882 new firms established in 24 cities of Yangtze River Delta in 2007, including 1498 private firms, 24 collective firms, 114 Hong Kong, Macao and Taiwan firms and 192 FDI, we examine how these characteristics affect the location choice of each ownership firms.

The results are consistent with our intuition. The regional market potential, land supply has a positive impact on private firms but not on FDI. However, air pollution as a proxy of congestion has a negative effect on both private firms and FDI. FDI has a tendency to locate near concentrated firms of the same industry. The constraints of finance affect the location choice of private firms to choose to locate in some cities.

In Chapter 3, we explore the factors impacting on the regional innovation capability of China at the prefecture level. By using an econometrics method, we examine how does the diversity of innovation, government expenditure for science and technology, market structure and other elements have an influence on the regional innovation capability.

In our analysis, we adopt a modified knowledge production function (Griliches, 1979; Buesa et al., 2010). Our dataset is a panel data that includes 286 prefecture cities of China in the period 2001-2008. Our analysis concerns about the role of diversity of regional innovation and whether the high-tech industrial zones play in promoting the innovation capability of cities.

We use the count of invention patent application in each city as a proxy of regional innovation capability to be the dependable variable. Because we utilize a count dataset, we assume that the data follows a negative binomial distribution. The independent variables include the diversity of regional diversity, high-tech industrial zones dummy, and control variables to control regional heterogeneity. Furthermore, in order to show the effect of the high-tech industrial park, we make an interaction term with the industrial park and the diversity of innovation.

The result shows that the diversity of regional innovation has a positive impact on regional innovation capability. On the contrary, the high-tech industrial zones have a negative impact on it. This can be interpreted that the government's policy is contrary to expectations and is a barrier to improving the region's innovation capability. Furthermore, when adding the interaction term, the effect of the high-tech industrial zone is no longer significant, and the interaction term presents a significant negative effect. This can be interpreted that it is better for the industry in the industrial park to gather similar industries rather than diverse industries.

In Chapter 4, we simulate the localization of innovation activity in the Yangtze River Delta by using a kernel density of a pair-wise firms' distance. The literature study the spatial distribution of innovation activity of China is using a provincial level data, existing the aggregation issues we mentioned before. Our study is based on microdata, hence we can provide an exact scope of localization.

By constructing counterfactual from the non-innovationes randomly chosen pairs of any firms in the same industry, including innovative enterprises and general enterprises, we can identify if the density of innovative firms are significantly closely located with each other than the firms in the same industry in general.

The dataset consists of 176,251 enterprises located in the Yangtze River Delta, of which 8584 enterprises carry out innovative activities. We classify these enterprises into three-digit classification to investigate the localization of innovation activities in each industries. The results show that not all industries active in innovation tend to concentrate. The interpretation of the results may be given as these industries are more inclined to take advantage of knowledge spillovers to promote their innovation activities.

In the last chapter, we conclude the findings of our study.

Chapter 2 The Location Choice of New Firms in Yangtze River Delta, China

2.1 Introduction

Since Reform and Opening-up, the economy of China has experienced rapid growth, especially in the eastern coastal areas. Due to the excellent geographic conditions that accessible to overseas market, the economic activities and population were gathering in the coastal regions. However, the reasons make some eastern coastal areas have rapid developments beyond the first-nature.

From the perspective of firms' location choice, the reasons for second-nature expand and sustains regional economic growth could be figured out. It is worth to explore what factors affect firms to choose their location in some specific regions. There exist lots of researches which show these determinants of firms make their location choice.

Firstly, firms usually make their location choice based on the profit maximization condition. They tend to locate in areas with high market potential. In the literature, the market potential is calculated as a sum of local and neighboring market size weighted by the distance (Harris, 1954). It was used to explain the factor of regions attract firms locating in there (Friedman et al., 1992; Disdier and Mayer, 2004; Head and Mayer, 2004). On the other hand, firms minimize their costs. In this situation, firms tend to locate in the regions where wage and land cost are lower.

Secondly, as Marshall (1890) pointed out, the externalities of agglomeration drive firms to choose the region with more specialized input, abundant skilled labor, and more chance to exchange information. In empirical researches, agglomeration economies were proved as an important factor that affects the location choice of firms (Diseider and Mayer, 2004; Crozet et al., 2004; Cheng and Stough, 2006; Head et al., 1995, 1999). Their studies show that the agglomeration effect has a positive influence on firms' location choice.

However, most of the studies on the determinants of location choices focus on the regional characteristics, but they ignore the individual characteristics of firms which also have influences on the decisions. It is understandable because their samples are the location choice of multinational enterprises (MNEs) mostly. The MNEs only need to consider the issue of maximizing profits or minimizing costs when making decisions to locate in foreign countries. But when the study of what kind of area is attractable for new firms, not only FDI but also lots of private enterprises, it is important to take into account individual characters of firms. Especially private enterprises often face financial constraints (Stein, 2003; Poncet et al., 2010).

This chapter, therefore, has two targets. One is exploring the regional characters, and the firms' individual characters affect the new firms in their location choices. The other one is to compare the difference in the location choice between private firms and foreign investments through capturing individual characteristics. It can fill the gap that lacks the study on private firms, although they are important parts of the regional economy.

By using the mixed logit model, we investigated the location choices of new firms located in the Yangtze River Delta (YRD) in 2007. In this year, there were 1848 manufacturing firms established in the Yangtze River Delta[1] [1](see Figure 2-1).

The estimation result shows that firms prefer to locate in a city with greater market potential, more land supply, and less congestion. The agglomeration effects affect the foreign investments' location choice positively while affecting private firms on both sides. The financial constraints affect some private firms when choosing a city to locate while not influence the location choice of foreign investments.

The rest of this chapter is organized as follows. Section 2 outlines the literature related to the location choice of firms and the hypotheses of our study. The econometric model is described in Section 3. Section 4 introduces the data. The estimated results are presented in Section 5. Section 6 presents the conclusion.

2.2 Literature reviews and hypothesis

The question of what factors affect new firms locate where they do is one of the most concerned topics in the spatial economy literature. Since Carlton (1979) first proposed an econometric model make it filled. Carlton's study originally introduced the discrete choice model (McFadden,1974) to analyze the location choice of new firms. Arauzo et al. (2010) state that using discrete choice models to analyze the

[1] The cities belong to YRD are: Shanghai, Nanjing, Wuxi, Xuzhou, Changzhou, Suzhou, Nantong, Lianyungang, Huaian, Yancheng, Yangzhou, Zhenjiang, Thaizhou, Suqian, Hangzhou, Ningbo, Wenzhou, Jiaxing, Huzhou, Shaoxing, Jinhua, Quzhou, Zhoushan, Taizhou, Lishui

location choice has an advantage on they can take into account both firm and regional factors.

2.2.1 Literature on industrial location choice

The new firms tend to choose the places where have a large market size to maximize their profits or have industrial agglomeration, inexpensive labor supply, and other factors to minimize their costs. When researchers pay their attention to the characteristics of regions, the conditional logit model is a most used model (Carlton, 1979; Head et al., 1995; Head and Mayer, 2004; Cheng and Stough, 2006). In some of the literature, the characteristics of firms are used as dummies for the choices to control the potential influence from these characteristics (Friedman et al., 1992; Levinson, 1996; Autant-Bernard, 2006).

Some studies address how the new firm characteristic affects their location choice as well. For instance, Arauzo and Manjón (2004) use a multinomial logit model to assess the firm size, and the industrial mix affect the new firms' location choices in Catalan. Their study shows that not only the regional characteristics but also the individual characteristics do influence on the location choice of some new firms.

One another discrete choice model often used to analyze the location choice of new firms is the nested logit model. In this model, similar alternatives can be partitioned into sets and allow the correlation between choices through nesting them. Comparing with the conditional (multinomial) logit model, it has an advantage on relaxing the independent of the irrelevant alternatives (IIA) assumption. Some specific factors such as historically specific ones are used as the division of nests (Hansen, 1987; Henderson et al., 1995; Henderson and Kuncoro, 1996). But how to define the nests is still a question in most cases that not suitable for division into nests.

In contrast, the mixed logit model is a more flexible alternative for including both regional characteristics and individual characteristics in the model. In recently, studies using a mixed logit model is increasing. For example, Rasciute and Pentecost (2008) use a mixed logit model examine the FDI' location choice in central and eastern European and find the probabilities of choices to choose a country varies both across sectors and different characteristics such as sizes of firms.

2.2.2 The literature on location choice in the case of China

The studies on industrial location choice in China is deeply influenced by the background of the times. At the beginning of 21 century, as FDI flourished in that time, a large number of studies on the location choice of FDI from the US, Japan, Taiwan or somewhere else arose. These studies mostly focus on the agglomeration effects influencing the location choice of FDI (Cheng and Kwan, 2000; He, 2002; Chang and Park, 2005, Du et al., 2008).

As time goes, labor cost was low in the past, but now it has risen even exceeded the benefits of economic agglomeration. The studies on the location choice of FDI in recent years is decreasing, and the focus has shifted into the determinants affecting the location choice of firms depending on the ownership. For instance, Zheng and Shi (2018) compare the effect of industrial land policy on the location choice of domestic firms with joint ventures and find that joint ventures are less intensive to the policy.

2.2.3 Research question and hypothesis

Although the determinants of location choice of domestic firms in China start getting attention, it still exists gaps to know more about the difference between domestic firms, especially the private firms and foreign investments. For example, private firms enjoy agglomeration effects, while foreign firms do not.

In addition, most of the previous works use a conditional logit model that emphasizes the regional characteristics as the factors influencing on the location choice of new firms, but they neglect the influence from the characteristics of firms' individual characteristics. In order to consider both factors simultaneously, we adopt a mixed logit model.

2.2.3.1 The roles of regional characteristics

Head and Mayer (2004) use a sample of Japanese plants location choice to show that "where the markets are" matters for their location choice. In some studies that about the case of China also have investigated the factor of market access affecting firms location choice (Amiti and Javorcik, 2008;

Cheng and Stough, 2006). Because the focus on our study is the comparison of determinants of location choice between private firms and foreign investments, the local market potential is an essential factor that cannot be neglected. In the Yangtze River Delta, unlike the foreign firms that chose location considering the accessibility to the export markets, the market for private firms is mainly domestic. Therefore, we set the following hypothesis:

Hypothesis 1. The local market potential is an essential factor affecting new private firms' location choice in a positive way, but it does not have a significant effect on foreign firms' location choice.

In China, the land is owned by the nation since the 1950s. The local governments have the rights to assigning land use by means of administrative allocation. As Zheng and Shi (2018) point out, industrial input increase will influence the location choice from many perspectives. For example, the land price can be lowered through increasing land supply. Therefore, our next hypothesis will be:

Hypothesis 2. In the context of China, an increase in land supply promotes the location choices of new firms. This effect does not depend on the type of ownership.

The traffic congestion is regarded as a force of dispersion for the location choice of new firms. The congestion not only increases the costs of transporting raw materials and products but also caused air pollution problems that endanger workers' health. Following Hou (2016) which found that local traffic congestion has a negative influence on the location choice of production-related activities, our hypothesis is:

Hypothesis 3. The congestion will have a negative influence on the location choice of new firms.

2.2.3.2 The roles of individual characteristics

The agglomeration effects are shown as influential factors that affect the location choice of new firms in the literature (Carlton, 1979; Head, 1995; Crozet et al., 2004; Disdier and Mayer, 2004). Although most previous studies adopt the number of firms belonging to the same industry as a measure of agglomeration, it is unsatisfactory for the case of large-scale firms. Ellison and Glaeser (1997) filled such a shortcoming by presenting a method to measure the geographic concentration by controlling for industrial characteristics if the geographic concentration is caused by an establishment of a large plant or by the agglomeration of a large number of small firms. Because we have an interest in how the agglomeration effect influencing the location choice of new firms, we adopt the Ellison-Glaeser Index (hereafter EG Index), which captures characteristics of concentration in different industries.

We follow the literature to assume that foreign investments are positively influenced by the agglomeration effects when choosing a city to locate. But it is not evident if private firms are influenced in the same way because they usually start with small/medium-scale and relatively low productivity, which make them hard to enter in big cities. Therefore, we hypothesize that,

Hypothesis 4. The foreign investments belong to industries with higher EG index entry to cities actively, but this may not be true for private firms.

The financial constraints are important factors for the location choice of new firms. As examined in many studies, financial constraints are an obstacle to firms' investments and growth (Hubbard, 1998; Poncet et al., 2010). Because of the large fixed costs, the single establishment firms (e.g., most of the private firms) are harder to relocate than the branch plants (e.g., most of the foreign investments) (Carlton, 1979). Thus, we consider the financial constraints of new private firms can have an influence on their location choice.

In addition to hindering the growth of firms, financial constraints may also reduce the firms' ability to withstand risks. However, foreign investments don't have to face these because they can obtain financial support from parent companies in the case of multinational enterprises or prioritized financial funding through preferential policies for foreign companies. Based on the different conditions between private firms and foreign investments, we will examine that,

Hypothesis 5. The financial constraints of private firms affect their location choice but have no influence in the case of foreign investments.

2.3 The Models

Arauzo et al. (2010) surveyed empirical studies on the methods of industrial locations. There are two approaches usually used when assessing what factors affect the location choice of firms. One is the discrete choice approach, and another is the count data approach. Because we use firm-level microdata, the discrete choice model allows us to analyze firms' location choice behavior in detail.

2.3.1 A Conditional Logit Model

The conditional logit model is the most used model in empirical studies on industrial location choice, which based on the random utility theory (McFadden, 1974). Profit for new firm i at location j is

$$\pi_{ij} = \beta' X_j + \varepsilon_{ij} \tag{1}$$

where X_j is a vector of characteristics of city j, β is a vector of estimated coefficients, and ε is a random error term.

A city j is selected by a new firm i if and only if:

$$\pi_{ij} > \pi_{ik},\ for\ j \neq k. \tag{2}$$

When the error terms are assumed to be independently and identically distributed, the probability that new firm i will locate in city j can be represented as:

$$P_{ij} = \frac{\exp(\beta' X_j)}{\sum_{k=1}^{K} \exp(\beta' X_k)} \tag{3}$$

where K is the number of cities in YRD. The estimation of parameters obtains from the maximum likelihood method.

The conditional logit model only contains the alternative-specific characteristics while in our case the characteristics of new firms also may affect their location choice. The mixed logit model includes both alternative-specific and case-specific attributes that are more flexible. It will be introduced in the next part.

2.3.2 A Mixed Logit Model

As we mentioned in the previous paragraph, the mixed logit model can include characteristics that vary across alternatives and individuals. McFadden and Train (2000) show that the mixed logit model provides diverse patterns of discrete choice. For the mixed logit model, profit for new firm i at location j is

$$\pi_{ij} = \beta' X_j + \gamma_j' Z_i + \varepsilon_{ij} \tag{4}$$

where X_j represents a vector of characteristics of city j (alternative-specific variables), β is a vector of estimated coefficients on X_j, Z_i is a vector of characteristics of new firm i (case-specific variables), and γ_j is a vector of estimated coefficients on Z_i . ε is a random error term independently and identically distributed.

The probability of choosing city j for new firm i is given by

$$P_{ij} = \frac{\exp(\beta' X_j + \gamma_j' Z_i)}{\sum_{k=1}^{K} \exp(\beta' X_k + \gamma_k' Z_i)} \quad . \tag{5}$$

The parameters are estimated by maximum simulated likelihood method.

2.4 Data and variables

Our sample consists of 1822 location choice of new firms in Yangtze River Delta (YRD) which set up their operation in 2007. The data comes from the Chinese Industrial Enterprises Database compiled by the National Bureau of Statistics of China. The ownership of new firms is classified as collective, private, foreign investment from Hong Kong, Macau, and Taiwan (HMT), foreign investment from other countries. We dropped 20 new state-owned enterprises which started up in YRD in this year because of their location choice mostly drive from an administrative instruction. Among the 1822 new firms, 24 firms are collectively owned, 1498 firms are private, 114 are owned by HMT investment, and 192 are owned by foreign investment.

There are 25 cities included in the Yangtze River Delta, namely Shanghai, Nanjing, Wuxi, Xuzhou, Changzhou, Suzhou, Nantong, Lianyungang, Huaian, Yancheng, Yangzhou, Zhenjiang, Thaizhou, Suqian, Hangzhou, Ningbo, Wenzhou, Jiaxing, Huzhou, Shaoxing, Jinhua, Quzhou, Zhoushan, Taizhou, Lishui. In our sample, Yancheng is dropped for it had no new firm this year (see Figure 2-2)

To find the difference between the determinants of private and foreign investments' location choices, we divide the sample of private firms only and foreign investments only. As only 17 cities are chosen by the foreign investments this year, the alternatives of location choice are reduced to this range when using the sample of foreign investments only. Complete description and resources for all variables are provided in Table 2-1.

Considering the problem of endogeneity, the data of independent variables and control variables are from 2006 data.

2.4.1 Independent Variables

Market potential. Firms tend to choose a location where the own market size is large market or having a good access to large markets. The previous studies such as Friedman et al. (1992) and Head and Mayer (2004) considered a spatial effect of nearby regions and measured the potential market size by using the Harris (1954) market potential function. Using the same approach, we calculate the market potential by the following formula:

$$MP_j = \sum_r S_r/(d_{rj})^{\rho}. \tag{6}$$

The market potential of region j is decided by the sum of the regional total retail sales S_j and other regions' total retail sales that weighted by inverse-distance. We use the parameter ρ to capturing the effect of distance decay. We give the value of ρ to 0.5, 1, 2 to see how the changing of sensitive on distance could make the effect of market potential varies on the location choice. In our models, it is an alternative-specific variable which varies with cities.

Land supply. The shortage of land supply will increase the land price and the production costs. Cheng and Stough (2006) take into consideration the land cost by using average house prices as an indicator. In our study, as the price of manufacturing land is unavailable at the level of the prefecture, we use the area of land requisition to indicate the land supply. This variable is an alternative-specific variable as well.

Air quality. The urban congestion effect is one of the elements that push for the dispersion of economic activities. In our study, urban congestion is indicated by using the annual average PM2.5 of cities. As vehicle emissions are considered to be the largest source of PM2.5 particles, the city with a high PM2.5 implies severe traffic congestion. The new firms tend to avoid urban congestion, so this will have a negative influence on the location choices of new firms. It is an alternative-specific variable, too.

Agglomeration effect. Yangtze River Delta is one of the most intense concentrations of industries in China. We have enough reasons to believe that new firms are motivated to locate in agglomerated cities. In the previous studies, the agglomeration phenomena are simply investigated by counting the number of industrial enterprises in the regions (Head and Ries, 1996; He, 2001). However, as scholars criticized, this measurement identifies agglomeration, including the situation where only one large-scale firm existing in the city, which does not match our perception of agglomeration. To remedy this problem, we calculate the Ellison-Glaeser index in 3-digit and in the city-level. The well-known equation is presented as follows:

$$\gamma = \frac{G^s - (1 - \sum_r (\lambda_r)^2) \times H}{(1 - \sum_r (\lambda_r)^2) \times (1 - H)}. \qquad (7)$$

Here G^s is the Krugman's Location Gini coefficient calculated as $G^s = \sum^R (\lambda_r - \lambda^s)^2$, where λ^s is the ratio of employment of s industry in the city r within the employment of industry s in the Yangtze River Delta. λ_r is the ratio of employment of all industries in the city r within the total industrial employment in Yangtze River Delta. H is the Herfindahl-Hirschman Index (HHI), namely $H = \sum_i (z_i)^2$, where z_i is the share of employment of i firm.

The Ellison-Glaeser Index is used as an indicator of industrial. It can show which cities are preferred for new firms when considering the degree of agglomeration of the industries they belong to. It is a case-specific variable that varying with individuals.

Financial constraints. Poncet et al. (2010) examined that private firms in China are credit constrained. The constraints are an obstacle to new firms because most of them are small and medium scale firms that have to mostly use up the cash to pay for workers, fixed assets, and so on. At the same time, they usually lack credit to loan from banks. These financial constraints may also have an influence on the location choice of new firms as weak to compete with the existing firms

In our study, two kinds of financial constraints are considered. The one is the financial constraint within new firms investigated as the ratio of each firm's cash flow to its total sales (Hall, 1992; Himmelberg and Petersen, 1994). The cash flow formula is:

$$cash\ flow = total\ sales - total\ industril\ intermediate\ input\ costs - VAT$$

$$+ accounts\ receivable - accounts\ payable + subsidy\ income \qquad (8)$$

Another one is the external finance. Like Cull et al. (2009) state, the key source of external funding in the case of China is loans from China. However, because of lacking the information on the quantity of loans in the Annual Industrial Enterprises Database, we need to construct a proxy for measuring the external finance. Follow Cai et al. (2005), the proxy is calculated as the ratio of interest payment to total sales by each new firm. The variables are case-specific and varies with individuals.

2.4.2 Control variables

Quality of local government. In China, as the local government plays important role in the socioeconomic activities, the quality of local government is a factor that cannot be dismissed when considering a firm location problem. Generally, firms prefer to locate in a city where the social environment is governed well so that their activities could be more efficient than these cities do not have good governance. Here we use two indicators to measure the quality of local government. One is the unemployment rate, and another is the wastewater treatment rate. They are varying with the alternatives.

Innovation capacity dummy. The new firms need to be competitive when trying to enter the market. In our study, firms with innovation capacity are identified by the dummy variable, which takes the value of 1 when a new firm enters the market producing a new product. This variable is case- specific variable varying with individuals.

Trade dummy. The firms having orders from overseas also could be regarded as competitive firms. We use a trade dummy to capture the competition of new firms. If the export delivery value of the new firm is not zero, the dummy value will be 1. And the value of dummy takes the value of 1. This variable is also a case-specific varying with individuals.

2.5 Empirical results

We present results for the location choice of all new firms, private firms only and foreign investments only separately in Table 2-2, Table 2-3, and Table 2-4. The coefficients of independent variables and control variables are estimated by maximum likelihood for both conditional logit regression and mixed logit regression. Comparing the magnitudes of coefficients, we are also concerned about the difference across the types of ownership.

2.5.1 The determinants of location choice in YRD

Table 2-2 shows the results of the conditional logit regression and the mixed logit regression for the entire new firms in Yangtze River Delta (except the state-owned enterprises). The Column (1), (2), and (3) provide the coefficients of alternative-specific variables without considering the characters of individuals and the two columns show the difference in the sensitiveness of distance to other markets.

Column (4), (5), and (6) provide the coefficients of both alternative-specific and case-specific variables. Results show us that the size of market potential has a positive and statistically significant influence on the location choices of entire new firms in YRD. This effect is sensitive to the distance to other cities. This is consistent with the theory of spatial economics in the literature (Head and Mayer, 2004; Cheng and Kwan, 2000)

The positive and statistically significant coefficient of land supply shows that new firms in YRD choose a city where has the land supplies in greater quantity. Similar results appear in Cheng and Stough (2006) which use house prices as an indicator of land supply.

The effect of urban congestion is negative and statistically significant, showing that urban congestion is a kind of dispersion forces that drives new firms to avoid choosing the city where it is congested.

Comparing with conditional logit regression, estimated coefficients of alternative-specific mixed logit regression are statistically significant and mostly have the same influence on the location choice of new firms.

Considering whether the industrial concentration affects the new firm in the same industry on their location choices, results vary across cities. The new firms are hard to make their location choice at Shanghai, Nanjing, Changzhou, Suzhou, Yangzhou, Zhenjiang, Thaizhou, Wenzhou, Shaoxing, Zhoushan as there is no space for new firms in same industries to enter. Conversely, the new firms are easy to choose a location at Lianyungang, Jiaxing, Huzhou, Jinhua, Taizhou utilizing the agglomeration economy.

The effects of financial constraints of new firms on their location choices are not very clear. When we include all types of ownership, its effect becomes mixed. We expect the effect will appear at the next regression that separated the ownership.

2.5.2 The difference in location choice between private firms and foreign investments

Table 2-3 and Table 2-4 separately presents the results for the location choices of private firms and foreign investments. Similarly, results include the estimation of both the conditional logit regression and the mixed logit regression.

The results of alternative-specific variables in the private firms' model are consistent with the full-sample estimate in the previous subsection. However, foreign firms exhibit a different tendency. For them, only the effect of urban congestion and the control of local government are statistically significant, and the signs of coefficients are consistent with the full-sample model and private firms' model.

The results of foreign investments have something interesting. We find that when only considerate the characteristics of alternatives, the market potential of regions has a significant and positive effect on the location choice of foreign investments. However, when we use a mixed logit model which controlled the effect of individual characteristics, the effect of regional market potential is significant at the 10% significance level when it is not sensitive to the distance. But if the market potential is sensitive to the distance, the local market potential is not significant to foreign investments' location decisions.

It may be explained that foreign investments are tending to locate at the place where agglomerated, these places are usually the same place with a big consumer market. After controlling the effect of industrial agglomeration, then we could find that the importance of local market potential is not very significant for foreign firms in China.

The difference in the influence of regional variables on the location choice of private firms and the foreign investments show that private firms face more external environment constraints than foreign investments when making a location choice. But the urban congestion equally affecting the firms with different ownership to make their location choices.

When comparing the difference in the influence of characters of individuals on their location choices, we find some interesting phenomena. The effect of industrial agglomeration has a different influence on the location choice of private firms and foreign investments. The foreign investments have a strong tendency to locate in the cities with a high concentration in the same industry, Suzhou, Nantong, Lianyungang, Jiaxing,Huzhou, most of these cities are close to Shanghai (except Lianyungang).

Thus, we can conclude that the private firms avoid locating in the cities with a high concentration in the same industries such as Shanghai, Nanjing, Wuxi, Changzhou, Suzhou, Yangzhou, Zhenjiang, Thaizhou, Wenzhou, Shaoxing. But some cities like Lianyungang, Jiaxing, and Jinhua are attractive in the industrial concentration for the new private firms in the same industries.

The effects of financial constraints are different in the two models as well. As we state in the previous section, unlike foreign investments without facing the financing constraints, the private firms face a difficult to financing. The cash flow and external financing variables are almost not statistically significant in the case of foreign investments. However, they have an influence on some private firms decision when choosing Thaizhou, Jiaxing, Jinhua, Quzhou to locate. The positive and statistically significant coefficients of external financing in these cities show that private firms with better financing ability are more likely to choose these cities to locate.

We also find a common point in the location choice of private firms and foreign investments. The innovative firms tend to locate in some cities belonging to Zhejiang Province, but this is not obvious in choosing cities belong to Jiangsu Province. The phenomena could be observed in both cases. It may be interpreted as the intensified competition in areas where entrepreneurship is active drives new firms to be innovative.

2.6 Conclusion

This chapter empirically examined the determinants of location choice of new firms starting-up their business in Yangtze River Delta in 2007. Our study mainly has some findings as follows.

For the influence of the characteristics of cities on the location choices of new firms, there are many differences between private firms and foreign investments. But there are still some common factors that affect their decisions.

Firstly, the local market potential is important to private firms but trifling to foreign investments as their market is mostly overseas. Secondly, it seems that private firms are more sensitive to the land cost than foreign investments as the land supply has an influence on the location choice of private firms, but foreign investments not. Thirdly, the urban congestion is a dispersion force that makes the new firms avoid locating in the city where is congested irrelevant to the ownership. Lastly, both private firms and foreign investments have a tendency to locate in the cities where the quality of local government is good.

For the influence of the characteristics of individuals (the new firms) on their location choice, we could find some differences as well.

Firstly, foreign investments are observed that take advantage of the benefits of agglomeration economies when making their location choice. They tend to locate in the cities with a high industrial concentration in the same industry. On the contrary, private firms can be observed that tend to choose some cities with a high industrial concentration but avoid some cities. Secondly, unlike foreign investments are not facing financing constraints, this factor has an effect on some cities to be selected or not by private firms. Thirdly, no matter foreign investments or private firms, the more innovative firms prefer to choose the cities in Zhejiang Province than these not. This may be interpreted as the intensified competition in areas where entrepreneurship is active drives new firms to be innovative.

This chapter confirmed the differences between private firms and foreign investments on the determinants of their location choices. While there are still some questions that needed to be studied, such as how about the TFP of new firms and does it affect the location choice. It is also possible that there exist some unobserved factors that affect the location choice such as spatial spillovers. Therefore, further research is needed to explore these questions.

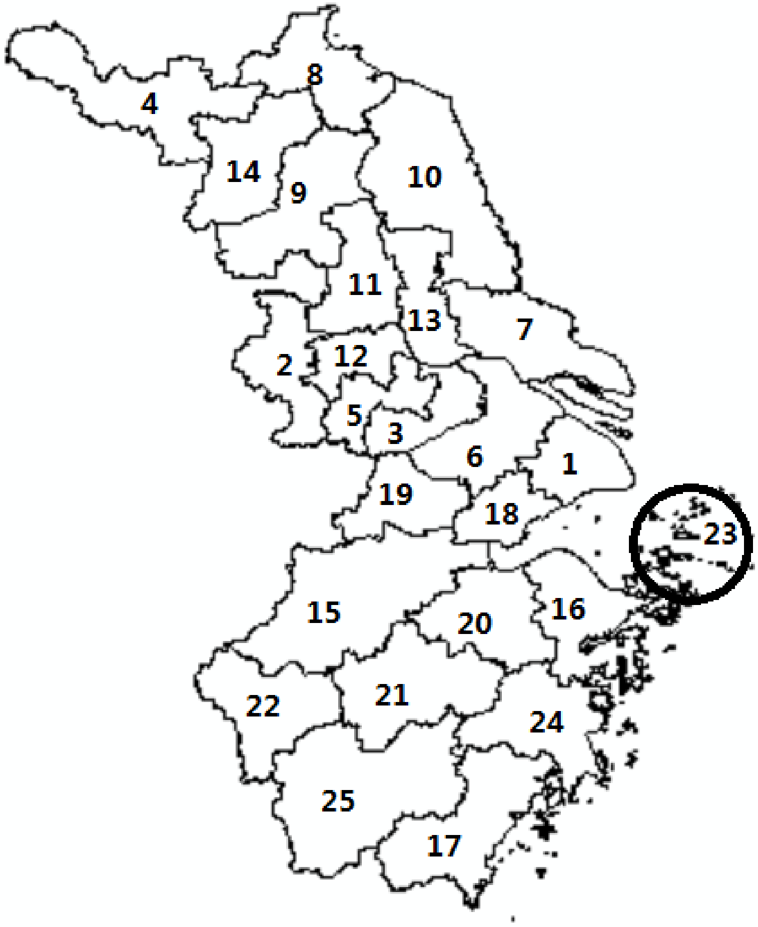


Figure 2-1. The distribution of cities in YRD[2]

[2] Numbers corresponding to the cities:1. Shanghai, 2.Nanjing, 3.Wuxi, 4.Xuzhou, 5.Changzhou, 6.Suzhou, 7.Nantong, 8.Lianyungang, 9.Huaian, 10.Yancheng, 11.Yangzhou, 12.Zhenjiang, 13.Thaizhou, 14.Suqian, 15.Hangzhou, 16.Ningbo, 17.Wenzhou, 18.Jiaxing, 19.Huzhou, 20.Shaoxing, 21.Jinhua, 22.Quzhou, 23.Zhoushan, 24.Taizhou, 25Lishui.

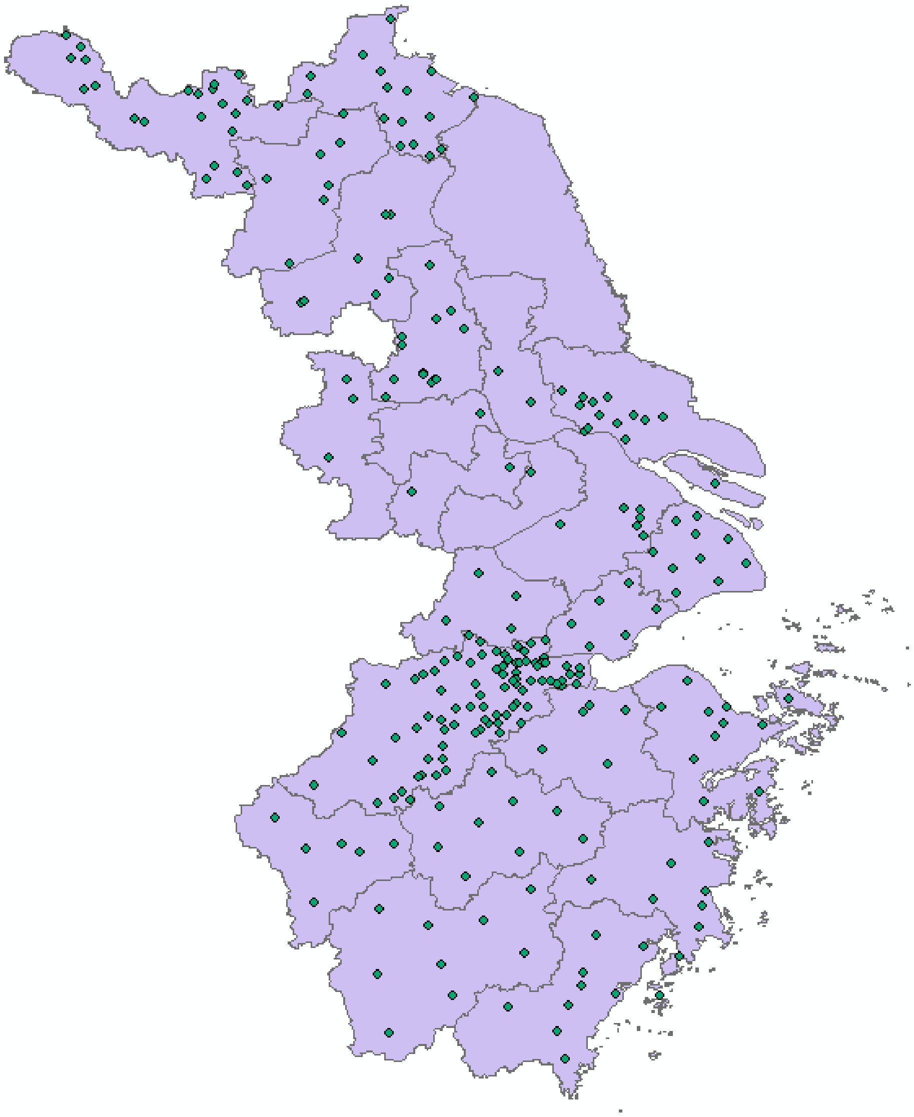


Figure 2-2. The location choice of new firms in YRD

Table2-1. Descriptive statistics

Variables	Definition	Source	unit	Mean	s.d.
MP	The sum of total retail sales from cities in YRD while discountig for distance	China Statistical Yearbook for regional economy(2006)	millon yuan	704.726	656.629
Land	The area of requisation of land in the city	China Urban-Rural Construction Statistical Yearbook(2006)	km^2	14.735	17.983
PM2.5	Annual average PM2.5 estimated value	Socioeconomic Data and Applications Center (2006)	$\mu g/m^3$	48.08	14.891
Unemployment	The unempolyment rate in the city	China Statistical Yearbook for regional economy(2006)	%	2.856	0.644
WTR	Wastewater treatment rate in the city	China Statistical Yearbook for regional economy(2006)	%	71.026	16.958
EG index	Ellison-Gleaser Index	Chinese Industrial Enterprise Database(2006)		0.044	0.041
Cash flow	The ratio of cash flow of new firms to the total sales	Chinese Industrial Enterprise Database(2007)	%	-0.214	6.888
EF	The ratio of external finance of new firms to the total sales	Chinese Industrial Enterprise Database(2007)	%	0.007	0.017
Innovation	The dummy of new firms produce new products or not	Chinese Industrial Enterprise Database(2007)		0.0159	0.365
Trade	The dummy of new firms export goods or not	Chinese Industrial Enterprise Database(2007)		0.255	0.436

Table2-2. Regression results (all new firms)

Explained variable	The location choice of new firms in YRD					
	CLM			MLM		
	(1)ρ=0.5	(2)ρ=1	(3)ρ=2	(4)ρ=0.5	(5)ρ=1	(6)ρ=2
Alternative-specific						
lnMP	0.814***	0.478***	0.282***	0.720***	0.456***	0.301***
	(0.137)	(0.063)	(0.046)	(0.257)	(0.141)	(0.105)
lnland	0.0715	-0.0158	0.115***	0.245***	0.162*	0.260***
	(0.044)	(0.047)	(0.038)	(0.082)	(0.096)	(0.076)
lnpm2.5	-2.228***	-2.186***	-2.132***	-1.321***	-1.264***	-1.197***
	(0.097)	(0.098)	(0.097)	(0.168)	(0.165)	(0.165)
unemployment	-0.181***	-0.112**	-0.124***	-0.258***	-0.230***	-0.229***
	(0.044)	(0.044)	(0.044)	(0.081)	(0.080)	(0.080)
WRT	0.0199***	0.0181***	0.0205***	0.0238***	0.0215***	0.0228***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
Case-specific						
Shanghai						
EG				-18.12***	-17.40***	-17.62***
				(4.851)	(4.737)	(4.773)
CF				0.082	0.0876	0.0867
				(0.189)	(0.192)	(0.191)
EF				-31.57**	-31.12**	-31.11**
				(14.620)	(14.670)	(14.640)
innovation				-0.396	-0.389	-0.397
				(0.557)	(0.557)	(0.557)
trade				-1.178***	-1.165***	-1.173***
				(0.396)	(0.396)	(0.396)
Nanjing						
EG				-75.79***	-80.04***	-79.61***
				(17.900)	(18.150)	(18.130)
CF				-0.081	-0.0856	-0.0849
				(0.318)	(0.310)	(0.311)
EF				-52.68	-55.68	-55.05
				(41.460)	(41.530)	(41.480)
innovation				-17.55	-17.56	-17.56
				(6633.300)	(6561.600)	(6540.300)
trade				-1.626	-1.682	-1.683
				(1.031)	(1.030)	(1.030)
Wuxi						
EG				-55.73***	-60.25***	-54.89***
				(16.160)	(16.780)	(15.990)
CF				-0.03	-0.0368	-0.0269
				(0.392)	(0.378)	(0.395)
EF				-73.32***	-75.83***	-73.32***
				(27.190)	(27.130)	(27.230)
innovation				-16.83	-16.84	-16.83
				(5239.000)	(5158.100)	(5213.900)
trade				-17.95	-17.96	-17.95
				(3803.300)	(3712.100)	(3799.600)
Xuzhou						
EG				5.599**	3.985	3.439
				(2.724)	(2.616)	(2.618)
CF				1.296***	1.187***	1.152***
				(0.325)	(0.323)	(0.325)
EF				-18.52	-21.04	-21.69

	(13.740)	(13.740)	(13.700)
innovation	-1.14	-1.147	-1.154
	(0.795)	(0.778)	(0.773)
trade	-1.855***	-1.893***	-1.906***
	(0.549)	(0.542)	(0.540)
Changzhou			
EG	-156.6***	-151.4***	-153.4***
	(39.510)	(39.620)	(39.590)
CF	-0.16	-0.157	-0.158
	(0.242)	(0.302)	(0.292)
EF	-119.2***	-117.4***	-118.0***
	(44.850)	(45.330)	(45.170)
innovation	-16.75	-16.77	-16.74
	(8046.300)	(8200.900)	(8011.400)
trade	-18.26	-18.22	-18.23
	(4854.300)	(4957.900)	(4882.500)
Suzhou			
EG	-23.60***	-24.13***	-22.83***
	(6.020)	(6.059)	(5.905)
CF	0.008	0.00725	0.0116
	(0.204)	(0.202)	(0.206)
EF	-47.84**	-48.69**	-47.53**
	(18.860)	(18.910)	(19.010)
innovation	-1.324	-1.337	-1.324
	(1.034)	(1.034)	(1.034)
trade	-0.698*	-0.713*	-0.688
	(0.424)	(0.424)	(0.424)
Nantong			
EG	-1.728	-1.915	-2.129
	(3.272)	(3.265)	(3.269)
CF	0.304	0.301	0.296
	(0.302)	(0.299)	(0.297)
EF	-21.72	-22.12	-22.28
	(15.160)	(15.210)	(15.190)
innovation	-0.409	-0.416	-0.424
	(0.629)	(0.629)	(0.629)
trade	-0.799**	-0.808**	-0.816**
	(0.399)	(0.399)	(0.399)
Lianyungang			
EG	14.07***	12.46***	13.04***
	(2.860)	(2.905)	(2.874)
CF	0.131	0.112	0.119
	(0.270)	(0.253)	(0.260)
EF	3.188	0.921	1.764
	(11.590)	(11.930)	(11.740)
innovation	-17.54	-17.53	-17.54
	(3765.500)	(3628.500)	(3693.600)
trade	-1.459**	-1.527**	-1.498**
	(0.618)	(0.616)	(0.617)
Huaian			
EG	-9.578	-10.57	-10.52
	(7.034)	(7.072)	(7.065)
CF	-0.001	-0.00361	-0.00307
	(0.331)	(0.326)	(0.327)
EF	-37.67	-39.07	-38.7
	(30.900)	(30.700)	(30.650)
innovation	-17.49	-17.5	-17.5
	(5755.200)	(5693.800)	(5687.500)

trade	-1.613	-1.643	-1.644
	(1.034)	(1.033)	(1.033)
Yangzhou			
EG	-14.57**	-12.28**	-14.85***
	(5.735)	(5.686)	(5.674)
CF	0.111	0.133	0.111
	(0.272)	(0.288)	(0.271)
EF	-11.09	-8.979	-11.34
	(15.350)	(15.120)	(15.340)
innovation	-1.115	-1.075	-1.129
	(1.036)	(1.037)	(1.036)
trade	-1.453**	-1.396**	-1.465**
	(0.612)	(0.614)	(0.612)
Zhengjiang			
EG	-105.1***	-98.04***	-101.7***
	(31.520)	(31.820)	(31.680)
CF	0.455	0.581	0.515
	(1.123)	(1.215)	(1.170)
EF	-80.15	-77.18	-78.63
	(53.970)	(55.340)	(54.660)
innovation	-16.89	-16.91	-16.9
	(7186.900)	(7334.200)	(7215.200)
trade	-18.11	-18.09	-18.1
	(4941.600)	(5093.400)	(4994.100)
Taizhou			
EG	-67.56***	-62.81***	-66.33***
	(18.210)	(18.170)	(18.200)
CF	0.586	0.673	0.61
	(0.696)	(0.736)	(0.708)
EF	18.89	19.55*	18.94
	(11.870)	(11.740)	(11.800)
innovation	-17.58	-17.61	-17.6
	(5656.000)	(5714.300)	(5666.700)
trade	-1.65	-1.586	-1.642
	(1.031)	(1.033)	(1.031)
Suqian			
EG	4.882	5.029	4.642
	(4.262)	(4.241)	(4.248)
CF	0.536	0.544	0.525
	(0.508)	(0.507)	(0.500)
EF	-46.04**	-46.36**	-46.39**
	(22.900)	(23.200)	(22.890)
innovation	2.060***	2.070***	2.044***
	(0.455)	(0.455)	(0.453)
trade	-2.572**	-2.575**	-2.580**
	(1.038)	(1.039)	(1.038)
Ningbo			
EG	0.841	0.46	0.511
	(1.841)	(1.841)	(1.839)
CF	0.144	0.138	0.139
	(0.115)	(0.114)	(0.114)
EF	4.701	4.317	4.319
	(5.197)	(5.208)	(5.187)
innovation	0.614**	0.601**	0.603**
	(0.266)	(0.266)	(0.266)
trade	0.410**	0.388**	0.391**
	(0.178)	(0.178)	(0.178)
Wenzhou			

EG	-7.830**	-9.399***	-9.007***
	(3.373)	(3.506)	(3.471)
CF	0.065	0.0563	0.0591
	(0.165)	(0.159)	(0.161)
EF	-30.78**	-33.02**	-32.28**
	(13.180)	(13.220)	(13.190)
innovation	-0.396	-0.438	-0.431
	(0.506)	(0.506)	(0.506)
trade	-0.824**	-0.875***	-0.865***
	(0.327)	(0.327)	(0.327)
Jiaxing			
EG	9.081***	9.194***	9.595***
	(2.077)	(2.063)	(2.054)
CF	-0.162**	-0.163**	-0.164**
	(0.074)	(0.074)	(0.074)
EF	5.533	5.676	5.948
	(5.684)	(5.677)	(5.641)
innovation	0.795***	0.803***	0.823***
	(0.295)	(0.295)	(0.295)
trade	0.373*	0.382*	0.408*
	(0.214)	(0.214)	(0.214)
Huzhou			
EG	3.488	4.219*	4.847**
	(2.525)	(2.478)	(2.468)
CF	0.057	0.0615	0.0656
	(0.144)	(0.146)	(0.149)
EF	-12.58	-11.51	-10.57
	(9.600)	(9.510)	(9.416)
innovation	1.437***	1.474***	1.505***
	(0.309)	(0.309)	(0.309)
trade	-0.568**	-0.541*	-0.514*
	(0.287)	(0.288)	(0.289)
Shaoxing			
EG	-19.84***	-19.14***	-18.81***
	(3.884)	(3.861)	(3.857)
CF	0.042	0.0428	0.0435
	(0.124)	(0.125)	(0.126)
EF	-0.274	0.00961	0.167
	(6.999)	(6.981)	(6.953)
innovation	1.802***	1.815***	1.820***
	(0.287)	(0.288)	(0.288)
trade	-0.073	-0.058	-0.0504
	(0.247)	(0.248)	(0.248)
Jinhua			
EG	4.958**	4.949**	4.606**
	(2.107)	(2.097)	(2.094)
CF	0.039	0.0389	0.0371
	(0.112)	(0.112)	(0.112)
EF	11.13**	11.14**	10.72**
	(5.284)	(5.283)	(5.264)
innovation	-1.048**	-1.046**	-1.057**
	(0.467)	(0.467)	(0.466)
trade	-0.529**	-0.528**	-0.542**
	(0.234)	(0.234)	(0.234)
Quzhou			
EG	-3.833	-3.537	-6.816*
	(3.591)	(3.566)	(3.922)
CF	-0.152*	-0.152*	-0.151*

			(0.078)	(0.078)	(0.079)
EF			14.63**	14.76**	13.07**
			(6.254)	(6.247)	(6.357)
innovation			-0.328	-0.321	-0.405
			(0.565)	(0.565)	(0.564)
trade			-0.851**	-0.842**	-0.954**
			(0.401)	(0.401)	(0.400)
Zhoushan					
EG			-23.02**	-18.99*	-19.94**
			(10.040)	(9.749)	(9.853)
CF			0.53	0.642	0.615
			(0.553)	(0.605)	(0.594)
EF			5.234	6.67	6.316
			(12.380)	(12.180)	(12.190)
innovation			1.611***	1.693***	1.666***
			(0.532)	(0.540)	(0.538)
trade			-18.29	-18.33	-18.32
			(2956.100)	(3032.300)	(3013.600)
Taizhou					
EG			3.644*	3.291*	3.229*
			(1.863)	(1.847)	(1.845)
CF			-0.077	-0.077	-0.0772
			(0.080)	(0.080)	(0.080)
EF			-4.909	-5.303	-5.383
			(5.821)	(5.833)	(5.817)
innovation			1.465***	1.449***	1.445***
			(0.248)	(0.247)	(0.247)
trade			-0.169	-0.185	-0.187
			(0.193)	(0.192)	(0.192)
Lishui					
EG			0.558	1.741	2.158
			(2.579)	(2.583)	(2.619)
CF			-0.167**	-0.168**	-0.169**
			(0.074)	(0.074)	(0.075)
EF			-12.44	-10.78	-10.09
			(9.661)	(9.551)	(9.493)
innovation			-0.176	-0.135	-0.12
			(0.447)	(0.447)	(0.448)
trade			-0.938***	-0.893***	-0.875***
			(0.324)	(0.325)	(0.326)
Log-likelihood	-5330.25	-5341.18	-4661.32	-4660.13	-4661.32
No. of firms	1821	1821	1820	1820	1820
No. of cities	24	24	24	24	24

Note: Conditional logit regressions and mixed logit regressions are all estimated by maximum likelihood. Standard errors in
paretheses.
Yancheng City is dropped in the regressions because of no new firms set up in there in 2007.
ρ is the parameter of market potential weighted by distance.
WRT is the wastewater treatment rate.

Table2-3. Regression results (private firms only)

Explained variable	The location choice of new private firms in YRD					
	CLM			MLM		
	(1)ρ=0.5	(2)ρ=1	(3)ρ=2	(4)ρ=0.5	(5)ρ=1	(6)ρ=2
Alternative-specific						
lnMP	0.512***	0.385***	0.235***	0.863***	0.622***	0.392***
	(0.154)	(0.070)	(0.051)	(0.291)	(0.159)	(0.115)
lnland	0.119**	0.0109	0.114***	0.188**	0.0394	0.188**
	(0.049)	(0.052)	(0.041)	(0.089)	(0.106)	(0.081)
lnpm2.5	-2.210***	-2.191***	-2.150***	-1.228***	-1.164***	-1.074***
	(0.107)	(0.107)	(0.107)	(0.184)	(0.182)	(0.183)
unemployment rate	-0.242***	-0.189***	-0.199***	-0.417***	-0.391***	-0.384***
	(0.048)	(0.049)	(0.048)	(0.093)	(0.092)	(0.092)
WRT	0.0197***	0.0172***	0.0190***	0.0265***	0.0233***	0.0251***
	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)
Case-specific						
Shanghai						
EG				-21.75***	-20.86***	-21.33***
				(6.293)	(6.140)	(6.204)
CF				-0.061	-0.0562	-0.0579
				(0.208)	(0.210)	(0.209)
EF				-18.63	-18	-18.23
				(18.130)	(18.070)	(18.060)
innovation				-0.303	-0.292	-0.307
				(0.640)	(0.640)	(0.640)
trade				-2.338**	-2.321**	-2.335**
				(1.025)	(1.025)	(1.025)
Nanjing						
EG				-73.71***	-80.19***	-79.16***
				(19.600)	(20.140)	(20.060)
CF				-0.075	-0.0801	-0.078
				(0.446)	(0.433)	(0.435)
EF				-171	-178.3	-176.2
				(125.300)	(125.100)	(124.900)
innovation				-17.33	-17.09	-16.58
				(6878.700)	(5970.800)	(4622.600)
trade				-17.93	-17.7	-17.2
				(5382.800)	(4625.000)	(3611.600)
Wuxi						
EG				-52.10***	-60.72***	-52.44***
				(16.080)	(17.220)	(15.980)
CF				0.07	0.0601	0.0722
				(0.453)	(0.433)	(0.452)
EF				-274.3*	-279.4*	-274.7*
				(160.700)	(156.600)	(160.500)
innovation				-17.26	-17.06	-16.54
				(6297.600)	(5494.000)	(4340.500)
trade				-17.86	-17.65	-17.12
				(5058.400)	(4338.400)	(3485.000)
Xuzhou						
EG				3.782	2.2	1.529
				(3.028)	(2.923)	(2.923)
CF				1.494***	1.272***	1.201***
				(0.466)	(0.444)	(0.439)
EF				-17	-20.17	-21.28
				(16.630)	(16.820)	(16.860)
innovation				-0.959	-1.015	-1.041

	(0.756)	(0.754)	(0.754)
trade	-1.835**	-1.888**	-1.910***
	(0.741)	(0.739)	(0.738)
Changzhou			
EG	-150.1***	-143.3***	-145.7***
	(40.140)	(40.150)	(40.110)
CF	-0.028	-0.0187	-0.02
	(0.622)	(0.634)	(0.630)
EF	-337	-335.4	-334.8
	(267.700)	(272.100)	(270.300)
innovation	-17.27	-17.04	-16.51
	(10492.700)	(9429.700)	(7170.100)
trade	-18.22	-17.91	-17.44
	(7044.800)	(6397.000)	(4917.300)
Suzhou			
EG	-75.24***	-78.61***	-74.16***
	(17.200)	(17.450)	(17.030)
CF	-0.12	-0.123	-0.116
	(0.366)	(0.359)	(0.368)
EF	-116	-120.3	-115.3
	(74.640)	(75.060)	(74.580)
innovation	-17.78	-17.53	-17.04
	(6536.700)	(5717.000)	(4505.200)
trade	-1.131	-1.167	-1.127
	(1.042)	(1.042)	(1.042)
Nantong			
EG	-4.249	-4.429	-4.746
	(3.884)	(3.880)	(3.887)
CF	0.404	0.394	0.386
	(0.433)	(0.428)	(0.424)
EF	-25.41	-25.98	-26.31
	(21.560)	(21.630)	(21.630)
innovation	-1.336	-1.343	-1.354
	(1.036)	(1.036)	(1.036)
trade	-0.995	-1.001	-1.013
	(0.620)	(0.620)	(0.620)
Lianyungang			
EG	11.66***	9.689***	10.48***
	(3.259)	(3.294)	(3.264)
CF	0.696	0.516	0.588
	(0.599)	(0.530)	(0.561)
EF	3.447	-0.207	1.313
	(16.440)	(16.960)	(16.630)
innovation	-17.74	-17.48	-16.99
	(4170.700)	(3522.700)	(2816.500)
trade	-0.736	-0.82	-0.784
	(0.633)	(0.630)	(0.631)
Huaian			
EG	-12.41	-13.61	-13.54
	(9.342)	(9.401)	(9.386)
CF	0.345	0.322	0.327
	(0.755)	(0.716)	(0.722)
EF	-272.5	-274.2	-273.4
	(174.400)	(171.900)	(172.300)
innovation	-16.65	-16.56	-16.09
	(4870.500)	(4583.700)	(3603.600)
trade	-17.37	-17.22	-16.74
	(4444.300)	(4070.200)	(3193.200)

Yangzhou			
EG	-14.89**	-10.91*	-14.65**
	(6.064)	(5.861)	(5.919)
CF	-0.019	0.00347	-0.0166
	(0.262)	(0.281)	(0.263)
EF	0.196	4.049	0.358
	(15.800)	(15.280)	(15.670)
innovation	-1.117	-1.035	-1.12
	(1.041)	(1.042)	(1.040)
trade	-1.18	-1.084	-1.178
	(0.747)	(0.750)	(0.747)
Zhenjiang			
EG	-108.8***	-98.05***	-103.5***
	(35.770)	(36.220)	(35.960)
CF	0.322	0.488	0.397
	(1.254)	(1.449)	(1.347)
EF	-89.67	-81.55	-85.38
	(102.400)	(100.700)	(101.500)
innovation	-17.36	-17.14	-16.63
	(9285.300)	(8441.700)	(6438.900)
trade	-18.03	-17.74	-17.26
	(7129.500)	(6574.200)	(4999.800)
Taizhou			
EG	-78.45***	-69.96***	-75.72***
	(22.640)	(22.460)	(22.520)
CF	0.48	0.626	0.525
	(0.874)	(0.959)	(0.901)
EF	27.87**	28.95**	27.89**
	(13.230)	(13.090)	(13.110)
innovation	-17.44	-17.21	-16.71
	(6138.900)	(5576.300)	(4252.700)
trade	-17.93	-17.66	-17.18
	(5150.600)	(4759.700)	(3585.500)
Suqian			
EG	5.729	6.225	5.605
	(4.396)	(4.369)	(4.378)
CF	0.488	0.516	0.482
	(0.610)	(0.628)	(0.606)
EF	-40.11	-38.95	-40.29
	(35.170)	(35.080)	(35.160)
innovation	1.807***	1.839***	1.800***
	(0.500)	(0.501)	(0.499)
trade	-18.16	-17.91	-17.42
	(3378.400)	(2992.900)	(2331.100)
Ningbo			
EG	-0.559	-0.998	-0.964
	(2.070)	(2.073)	(2.070)
CF	0.265	0.251	0.253
	(0.186)	(0.182)	(0.183)
EF	8.161	7.57	7.481
	(7.611)	(7.652)	(7.595)
innovation	0.38	0.364	0.366
	(0.306)	(0.306)	(0.306)
trade	0.642***	0.621***	0.623***
	(0.229)	(0.228)	(0.228)
Wenzhou			
EG	-6.991**	-8.863**	-8.380**
	(3.413)	(3.556)	(3.513)

CF	0.026	0.0129	0.0169
	(0.195)	(0.188)	(0.190)
EF	-33.35*	-37.47**	-36.13**
	(17.270)	(17.650)	(17.500)
innovation	-0.692	-0.744	-0.734
	(0.561)	(0.561)	(0.561)
trade	-0.241	-0.298	-0.286
	(0.356)	(0.356)	(0.356)
Jiaxing			
EG	7.996***	8.064***	8.648***
	(2.252)	(2.240)	(2.229)
CF	-0.218**	-0.219**	-0.222**
	(0.106)	(0.106)	(0.106)
EF	13.60*	13.83*	14.32*
	(7.954)	(7.964)	(7.850)
innovation	0.495	0.499	0.526
	(0.334)	(0.334)	(0.334)
trade	0.545**	0.547**	0.579**
	(0.260)	(0.260)	(0.261)
Huzhou			
EG	1.865	2.732	3.657
	(2.754)	(2.699)	(2.682)
CF	-0.106	-0.105	-0.104
	(0.136)	(0.138)	(0.139)
EF	-9.864	-8.222	-6.451
	(12.650)	(12.520)	(12.320)
innovation	1.007***	1.046***	1.090***
	(0.353)	(0.353)	(0.354)
trade	-0.612	-0.586	-0.556
	(0.382)	(0.383)	(0.384)
Shaoxing			
EG	-18.96***	-18.06***	-17.64***
	(4.224)	(4.192)	(4.181)
CF	0.141	0.151	0.157
	(0.229)	(0.233)	(0.236)
EF	-3.044	-2.269	-1.908
	(11.390)	(11.350)	(11.260)
innovation	1.313***	1.332***	1.338***
	(0.339)	(0.339)	(0.339)
trade	-0.021	-0.000379	0.00731
	(0.335)	(0.336)	(0.336)
Jinhua			
EG	3.527	3.703*	3.203
	(2.244)	(2.233)	(2.229)
CF	-0.003	-0.00229	-0.0052
	(0.141)	(0.141)	(0.141)
EF	16.49**	16.80**	15.85**
	(7.223)	(7.236)	(7.179)
innovation	-1.269**	-1.261**	-1.277**
	(0.508)	(0.508)	(0.508)
trade	-0.121	-0.114	-0.133
	(0.269)	(0.270)	(0.269)
Quzhou			
EG	-5.025	-4.756	-8.999**
	(3.741)	(3.725)	(4.134)
CF	-0.234**	-0.235**	-0.234**
	(0.106)	(0.106)	(0.106)
EF	24.54***	24.85***	21.96***

				(7.831)	(7.848)	(7.867)
innovation				-0.452	-0.446	-0.558
				(0.573)	(0.573)	(0.572)
trade				-0.246	-0.237	-0.375
				(0.417)	(0.417)	(0.416)
Zhoushan						
EG				-16.06*	-11.17	-12.72
				(9.380)	(8.906)	(9.070)
CF				0.876	1.171	1.074
				(0.745)	(0.809)	(0.792)
EF				16.02	18.18	17.21
				(12.890)	(12.720)	(12.670)
innovation				1.674***	1.816***	1.763***
				(0.553)	(0.566)	(0.561)
trade				-18.13	-17.88	-17.39
				(4095.600)	(3781.200)	(2922.500)
Taizhou						
EG				3.317*	3.054	2.913
				(1.991)	(1.974)	(1.970)
CF				-0.091	-0.0916	-0.0926
				(0.111)	(0.111)	(0.111)
EF				7.364	7.075	6.68
				(7.370)	(7.385)	(7.335)
innovation				1.384***	1.371***	1.365***
				(0.265)	(0.265)	(0.265)
trade				0.204	0.192	0.187
				(0.232)	(0.231)	(0.231)
Lishui						
EG				-0.312	1.328	1.799
				(2.727)	(2.734)	(2.781)
CF				-0.245**	-0.247**	-0.247**
				(0.105)	(0.105)	(0.105)
EF				-3.078	-0.385	0.448
				(10.930)	(10.800)	(10.710)
innovation				-0.312	-0.255	-0.236
				(0.462)	(0.463)	(0.463)
trade				-0.41	-0.353	-0.336
				(0.353)	(0.355)	(0.356)
Log-likelihood	-4347.79	-4347.79	-4352.71	-3791.33	-3787.84	-3789.91
No. of firms	1492	1492	1492	1492	1492	1492
No. of cities	24	24	24	24	24	24

Note: Conditional logit regressions and mixed logit regressions are all estimated by maximum likelihood. Standard errors in paretheses.

Yancheng City is dropped in the regressions because of no new firms set up in there in 2007. ρ is the parameter of market potential weighted by distance.

WRT is the wastewater treatment rate.

Table 2-4. Regression results (foreign investments only)

Explained variable	The location choice of new foreign investments in YRD					
	CLM			MLM		
	(1)ρ=0.5	(2)ρ=1	(3)ρ=2	(4)ρ=0.5	(5)ρ=1	(6)ρ=2
Alternative-specific						
lnMP	3.018***	1.211***	0.835***	1.844*	0.213	-0.0988
	(0.546)	(0.301)	(0.249)	(1.033)	(0.590)	(0.519)
lnland	-0.665***	-0.472**	-0.302	-0.156	0.276	0.457
	(0.195)	(0.201)	(0.184)	(0.406)	(0.423)	(0.412)
lnpm2.5	-2.887***	-2.259***	-2.129***	-2.192***	-1.685**	-1.661**
	(0.380)	(0.325)	(0.315)	(0.767)	(0.684)	(0.680)
unemployment rate	-0.206	-0.176	-0.161	0.055	0.111	0.141
	(0.149)	(0.142)	(0.143)	(0.343)	(0.342)	(0.347)
WRT	0.0190**	0.0177*	0.0220**	0.0423**	0.0486**	0.0537***
	(0.009)	(0.010)	(0.010)	(0.018)	(0.019)	(0.020)
Case-specific						
Shanghai						
EG				6.888	8.199	8.748
				(10.430)	(10.320)	(10.320)
CF				0.025	0.0576	0.0704
				(0.573)	(0.599)	(0.607)
EF				-44.31	-41.87	-40.95
				(31.430)	(31.360)	(31.310)
innovation				0.14	0.163	0.178
				(1.218)	(1.218)	(1.219)
trade				-1.417**	-1.345**	-1.309**
				(0.647)	(0.655)	(0.658)
Nanjing						
EG				-44.05	-47.79	-46.21
				(41.500)	(41.630)	(41.620)
CF				-0.126	-0.145	-0.136
				(0.742)	(0.720)	(0.733)
EF				-8.836	-11.29	-10.22
				(42.590)	(43.070)	(42.980)
innovation				-19.53	-19.52	-18.41
				(44718.300)	(43811.000)	(25162.000)
trade				-1.151	-1.25	-1.204
				(1.158)	(1.150)	(1.156)
Xuzhou						
EG				16.56	9.481	8.6
				(13.050)	(13.440)	(13.520)
CF				1.953*	1.45	1.395
				(1.077)	(0.996)	(0.986)
EF				-163.7	-161.2	-160.9
				(144.400)	(129.400)	(127.800)
innovation				-21.95	-21.38	-20.21
				(26360.800)	(28287.900)	(15954.900)
trade				-1.185	-1.424	-1.442
				(1.092)	(0.989)	(0.979)
Suzhou						
EG				19.66**	21.22***	22.16***
				(7.999)	(8.081)	(8.128)
CF				0.048	0.118	0.156
				(0.570)	(0.604)	(0.617)

EF	-27.86	-23.95	-22.03
	(23.750)	(23.150)	(22.800)
innovation	0.1	0.169	0.21
	(1.214)	(1.216)	(1.218)
trade	-1.332**	-1.205*	-1.132*
	(0.612)	(0.629)	(0.636)
Nantong			
EG	24.31***	22.81***	22.63**
	(8.831)	(8.793)	(8.819)
CF	-0.295	-0.307	-0.309
	(0.542)	(0.514)	(0.508)
EF	-42.87	-46.57	-47.37
	(38.910)	(39.340)	(39.420)
innovation	1.02	0.899	0.875
	(1.246)	(1.237)	(1.236)
trade	-1.448*	-1.527*	-1.533*
	(0.832)	(0.807)	(0.801)
Lianyungang			
EG	32.88**	32.57**	33.51**
	(13.450)	(13.530)	(13.610)
CF	0.615	0.6	0.649
	(2.123)	(2.150)	(2.228)
EF	-308.4	-301.6	-303.5
	(359.600)	(354.100)	(356.500)
innovation	-16.69	-15.5	-15.16
	(12935.500)	(7121.800)	(6116.400)
trade	-21.13	-20.17	-19.84
	(11182.700)	(7178.500)	(6176.800)
Huaian			
EG	-155.9	-151.2	-150.2
	(145.100)	(138.000)	(135.800)
CF	0.252	0.196	0.177
	(0.801)	(0.788)	(0.784)
EF	65.64	57.01	54.17
	(64.150)	(61.890)	(61.290)
innovation	-18.87	-18.36	-17.15
	(38472.100)	(36189.900)	(21111.600)
trade	-20.9	-20.76	-19.66
	(25488.700)	(23860.100)	(13761.200)
Yangzhou			
EG	4.789	-10.08	-18.64
	(27.290)	(32.910)	(35.710)
CF	2.629	1.725	1.419
	(1.922)	(1.793)	(1.690)
EF	-412.2	-410.6	-421.1
	(524.500)	(537.400)	(540.000)
innovation	-13.3	-15.31	-12.95
	(1954.100)	(5599.900)	(1661.400)
trade	-17.83	-19.94	-17.72
	(1963.700)	(5436.700)	(1654.200)
Taizhou			
EG	-8.87	-20.16	-27.29
	(30.730)	(34.140)	(35.730)
CF	0.427	0.185	0.0851
	(1.425)	(1.248)	(1.135)
EF	-18.22	-30.26	-36.81
	(58.510)	(62.370)	(63.860)
innovation	-19.59	-19.57	-18.46

	(44065.300)	(43093.900)	(24301.500)
trade	-0.777	-1.154	-1.324
	(1.294)	(1.243)	(1.217)
Suqian			
EG	5.615	0.593	-1.241
	(59.760)	(58.630)	(58.390)
CF	2.642	2.51	2.458
	(2.764)	(2.617)	(2.571)
EF	-1424.7	-1372.8	-1355.1
	(959.200)	(936.000)	(929.600)
innovation	-44.32	-42.61	-42.14
	(32705.100)	(19439.200)	(12003.800)
trade	-31.37	-30.01	-28.8
	(6545.900)	(4199.300)	(2500.000)
Ningbo			
EG	9.642	9.49	9.736
	(7.475)	(7.511)	(7.546)
CF	-0.213	-0.211	-0.209
	(0.274)	(0.276)	(0.276)
EF	3.591	3.372	3.521
	(9.635)	(9.655)	(9.677)
innovation	1.847***	1.834***	1.838***
	(0.702)	(0.700)	(0.701)
trade	-0.169	-0.183	-0.163
	(0.421)	(0.422)	(0.423)
Jiaxing			
EG	16.68**	17.62**	18.07**
	(8.374)	(8.416)	(8.454)
CF	-0.001	0.0102	0.0141
	(0.377)	(0.383)	(0.385)
EF	4.385	5.333	5.686
	(11.700)	(11.650)	(11.660)
innovation	1.411*	1.468*	1.483*
	(0.842)	(0.843)	(0.844)
trade	0.361	0.535	0.593
	(0.615)	(0.628)	(0.629)
Huzhou			
EG	18.86**	20.88**	21.21**
	(9.054)	(8.988)	(9.008)
CF	0.569	0.619	0.625
	(0.459)	(0.478)	(0.479)
EF	-32.08	-29.94	-29.72
	(25.640)	(25.600)	(25.590)
innovation	2.858***	2.991***	3.001***
	(0.818)	(0.822)	(0.822)
trade	-0.853	-0.705	-0.684
	(0.662)	(0.675)	(0.674)
Shaoxing			
EG	-25.08	-24.34	-24.79
	(16.060)	(16.060)	(16.110)
CF	-0.444	-0.445	-0.444
	(0.291)	(0.294)	(0.293)
EF	-5.771	-5.324	-5.509
	(13.910)	(13.870)	(13.910)
innovation	2.920***	2.920***	2.907***
	(0.793)	(0.792)	(0.790)
trade	-0.284	-0.233	-0.25
	(0.553)	(0.555)	(0.554)

Jinhua						
EG				9.28	7.494	7.44
				(12.290)	(12.420)	(12.480)
CF				0.027	0.0189	0.0191
				(0.408)	(0.404)	(0.404)
EF				19.43	18.53	18.54
				(12.190)	(12.240)	(12.270)
innovation				0.415	0.339	0.332
				(1.303)	(1.299)	(1.299)
trade				-0.971	-1.077	-1.076
				(0.825)	(0.805)	(0.803)
Taizhou						
EG				-0.165	-1.275	-1.383
				(13.040)	(13.200)	(13.260)
CF				-0.671**	-0.669**	-0.668**
				(0.300)	(0.302)	(0.302)
EF				-93.47**	-94.15**	-94.24**
				(39.710)	(39.610)	(39.610)
innovation				1.125	1.098	1.094
				(1.027)	(1.023)	(1.022)
trade				-0.159	-0.243	-0.248
				(0.659)	(0.652)	(0.651)
Log-likelihood	-474.88	-484.61	-488.24	-387.63	-389.26	-389.31
No. of firms	191	191	191	191	191	191
No. of cities	17	17	17	17	17	17

Note: Conditional logit regressions and mixed logit regressions are all estimated by maximum likelihood.
Standard errors in paretheses.
Only the cities have new foreign investments in 2007 are in the regression.
ρ is the parameter of market potential weighted by distance.

Chapter 3 The Determinants of Regional Innovative Capability in China: in a Prefectural Level

3.1 Introduction

As a driving force for urban development, the innovative capability of regions has attracted great attention in the literature. Researchers and policymakers are interested in the determinants of regional innovative capability that stimulate the growth of regional innovative activities. Through observing the distribution of innovative activities, it is found that these activities are concentrated in some regions.

In the 1990s, the New Economic Geography Theory arose and "the love of variety" is used to summarize the reasons of a concentration of people, firms, public service and innovation (Fujita et al.1999, Fujita et al.,2018). Consumers agglomerate into a place where the variety of consumption is greater and consumers can achieve higher utility under a given budget constraint. The firms also enjoy the variety of input goods to upgrade their productivities. Similarly, the utility of residents will be higher by enjoying a greater variety of public service. The situation of innovative activity follows suit. The region where the variety of innovators is greater could promote the region to be more creative.

This phenomenon also appears in China, where the economic and innovative activity is concentrating in the eastern region after the Reform and Opening-up. Based on the data from the State Intellectual Property Office of the People's Republic of China (SIPO), the ratio of the patent applications in the eastern provinces to the total had increased from 66% in 2001 to 74.4% in 2008. The innovative activity is concentrated in the eastern region of China. The eastern region of China has an advantage in natural conditions. However, the second nature plays a more important role in developing innovative creativities. It is worth exploring the determinants of the innovative capability of regions in China.

The most frequently mentioned determinants of innovative capability of regions are R&D expenditures (Jaffe 1989, Feldman and Florida 1994, Audretsch and Feldman 1996, Crescenzi and Rodríguez-Pose 2011), variety of regional knowledge (Jacobs 1969, Gleaser et al. 1992, Audretsch and Feldman 2004, Frenken et al. 2007), human capital such as people work in the universities or R&D institutes (OECD 1995, Amable and Petit 2001), the market structure (Cohen and Levin 1989, Carree and Thurik 1999, Buesa et al. 2010) and so on.

China has a strong power to control its economy activities by administrative methods. The administrative policies play important roles in affecting the innovative capability of regions. The most characterized policy to stimulate regional innovation is the establishment of the National High-tech Industrial Development Zone (NHIDZ) in China. Because it is designed to achieve the strategy of developing national high-tech industries and commercialize research results smoothly.

In the existing literature, due to the constraint of available data and the actual needs of attracting foreign investment, most research on the regional innovation distribution and its determinants are analysis at a provincial level or focus on the influence of FDI. For instance, Liu and Sun (2009) use methods such as rank-frequency to compare the spatial distribution of innovative activities in China and the U.S.

The main findings in the literature are the rapid growth of invention patents in China and the concentration of innovative activities from inland areas to coastal areas. These studies are mostly based on provincial data, which is a wide range, but innovative activities are concentrated in a narrow range in the prefectural level.

Recently, studies on the prefectural innovation creativities are emerging. For instance, Tan et al. (2017) using a panel dataset of the granted patent in 336 cities examined the spatial impact on the regional innovation by using Gini coefficient, Moran' I and spatial Durbin model. Their research showed that the invention patent is mainly concentrated in large coastal cities, and the spatial autocorrelation of the invention patent is insignificant.

Thanks to the recently available patent database which removes the limitation of data, it is possible to explore the determinants of regional innovation at prefecture-level from many different angles and using new methods to learn more about the situation of innovation in China.

This chapter aims to explore the determinants of the innovative capability of regions in China. We also examine whether the administrative policy plays the function of enhancing the innovative capability of regions or not. Through using the interaction term of the variety of innovation and NHIDZ, we estimated the effectiveness of the policy. This chapter contributes to examining the policy effect of establishing a high-tech zone in China.

Using the panel data on Chinese invention patent applications at a prefectural level including 286 prefectures during 2001-2008, this chapter will show that variety of innovation and high-tech zone have a positive effect on regional innovative capability, but those effects are over-estimated as we can see that their interaction term have negative coefficients.

The rest of this chapter is organized as follows. Section 2 introduces the NHIDZ and develops five hypotheses on the determinants of the innovative capability of regions. The model and measurements of variables are presented in Section 3. Section 4 presents data, empirical method, and results. The conclusion is presented in Section 5.

3.2 The determinants of the innovative capability of regions

3.2.1 The role of NHIDZ in the innovative capability of regions

The National Hi-Tech Industrial Development Zone (NHIDZ) is defined as a specific industrial zone that is concentrating on promoting the development of high-tech industries. In 1991, China State Council approved 26 prefectures to set up NHIDZ. Then more 25 prefectures were added in 1992. The total number of NHIDZ was kept at 53 for ten years after Shannxi Yangling Agriculture Hi-Tech Industrial Development Zone was approved in 1997.

Being different from these industrial zones are called “China National Economic and Technological Development Zones” which aims at increasing economic growth and attract foreign direct investments (FDI), NHIDZ focus on encouraging local firms to produce new products and to be innovative.

The most fundamental task of NHIDZ is developing high-tech industries with a nationally created intellectual property. To achieve this task, some preferential policies are formulated for high-tech firms in costs production, increasing returns, and other aspects.

For the sake of reducing production costs, the preferential policies include administrative procedures, tax policy, and financial support. For example, high-tech enterprises registered at NHIDZ are approved to simplify the procedure of obtaining a license to import raw materials and parts from overseas. The tariff will be exempt when exporting high-tech products. Banks provide loans and bonds services to help high-tech enterprises’ financing.

To support firms to increase profit, the preferential policies include independent pricing, accelerated depreciation on equipment, and exemption from the obligation to purchase government bonds. In China, commodity prices are subject to NDRC’s recommendation. However, enterprises located in the zones benefit from preferential approval for their new product to stimulate enterprises' innovation. Through reducing the burden on high-tech enterprises, it is expected to attract more high-tech enterprises to aggregate in the zones.

Besides, other preferential policies to the zones include a tax credit for the construction of zones. Or senior personnel will be given priority to be employed in the zones.

Because of these functions of NHIDZ, the high-tech industries developed in the high-tech zones. It is the original intention of the Chinese State Council to establish the NHIDZ. The effectiveness of high-tech industrial zones has been discussed in many pieces of literature.

Some empirical analysis shows that the industrial zone has a small effect on promoting the commercialize of research results or enhancing innovation. For instance, Felsenstein (1994) examined the relationship between innovation and science park and found that the science park has effects on entrenching innovation rather than inducing it. But there also exists some literature which confirms that the industrial zone has a positive influence on the firms' innovation. Albahari et al. (2013) show that firms located in the S&T park have a better innovation performance.

In the context of China, Lai et al. (2005) explored the effects of the industrial cluster and regional economic policy on innovation capacity by comparing the developed cluster in Shanghai, Kunshan, Shenzhen, and Dongguan. Fan (2003) evaluated the technology innovation capacity of 52 national high-tech development industrial zones in China.

The existing literature sheds light on the innovation performance of firms located in the high-tech industrial zone to examine the relationship between the industrial zone and regional innovation capability. It is not appropriate to use the innovative performance of companies within an industrial park to represent the innovative capability of the entire region. Therefore, in this chapter, this relationship will be examined by using the number of annual patent applications as a proxy of regional innovative capability.

To identify the effect of NHIDZ on the regional innovative capability more precisely, the dataset includes the information regarding whether each prefecture owns an NHIDZ or not. It is expected that NHIDZ will have a positive effect on promoting the innovative capability of prefectures. The hypothesis is as follows.

Hypothesis 1: The prefecture which has an NHIDZ will have a higher innovative capability.

3.2.2 Other determinants of the innovative capability of prefectures

3.2.2.1 The variety of innovative activities

The diversity of knowledge plays a vital role in fostering the innovative capability is generally argued in the literature (Jacobs 1969, Duranton and Puga 2001, Berliant and Fujita 2011). The more differentiated new knowledge exists in the region, the easier to create a new idea through getting inspiration from different types of knowledge. The variety of innovative activities in a region will generate knowledge spillover to fostering the regional innovative capability. The hypothesis is given as follows:

Hypothesis 2: The higher of the variety of innovation in one prefecture, the higher innovative capability of this prefecture.

3.2.2.2 The innovative inputs

The knowledge production function, originally developed by Griliches (1979), shows that the generation of a new product upon innovative inputs such as R&D expenditure and human capital. The R&D expenditure and the number of educated students are essential factors on the regional innovative capability. Sufficient resources for innovation are necessary conditions for generating plenty of innovation output. Thus， two hypotheses are given as follows:

Hypothesis 3: The more R&D expenditure costs in the prefecture, the higher innovative capability of the prefecture will be.

Hypothesis 4: The greater number of educated student residents in the prefecture, the higher innovative capability of the prefecture willbe.

3.2.2.3 The market construction

Chinese domestic market consists of firms of five types of ownership and their mixture. Such a market structure also affects the innovative capability of prefectures. A region with a higher weight of state-owned enterprises may have less marketization. As Nie et al. (2016) pointed out, the state-owned enterprises and collective enterprises have a high proportion of zombie company which should have gone bankrupt but still operated by the subsidiary of government. It is a disadvantage for generating innovation in a low marketization region. Accordingly, the hypothesis is given as follows:

Hypothesis 5: Prefectures with a higher level of marketization will have a higher innovative capability.

Considering the knowledge spillover from FDI as being pointed out in the literature, the market with a greater presence of foreign enterprises will improve the innovative capability of this prefecture. Then the hypothesis is given as follows:

Hypothesis 6: The higher the influence of FDI in the prefecture, the higher innovative capability the prefecture will be.

3.3 The model and measurement of variables

3.3.1 The model

The most commonly used empirical approach to discuss the determinants of regional innovation capacity is the knowledge production function that was initially developed by Griliches (1979). In this chapter, a modified knowledge production function proposed by Buesa et al. (2010) is adopted. The model is defined as follows:

$$Patent_{r,t} = \alpha_0 + \alpha_1 RD_{r,t-1} + \alpha_2 HC_{r,t-1} + MS_{r,t-1}\eta + \beta HZ_r + \gamma VRT_{r,t-1} + \varepsilon_{r,t} \quad (9)$$

where $Patent_{r,t}$ is the number of firms' patents applications in prefecture r at time t, $RD_{r,t-1}$ is the log of R&D expenditure of prefecture r at time $t-1$, $HC_{r,t-1}$ is the human capital given by the number of college students in prefecture r at time $t-1$, $MS_{r,t-1}$ is the market structure of prefectures expressed by the degree of marketization and globalization in prefecture r at time $t-1$, and HZ_r is a dummy variable representing whether there is a national high-tech industrial development zone in prefecture r. The latter takes the value 1 if the prefecture has an NHIDZ and 0 otherwise. The number of NHIDZ was almost unchanged during the period of this chapter. The number had remained at 53 (include the four municipalities) until Ningbo City received approval to set up a new one. Thus, HZ_r is treated as invariant with time. $VRT_{r,t-1}$ is the variety of innovation in prefecture r at time $t-1$. The last term $\varepsilon_{r,t}$ is an idiosyncratic error.

Considering that there is a time lag between the development and application of patents, the time of information of prefectures are using the year before application years to control it.

Because the dependent variable is a count data, a Poisson regression or a Negative Binomial regression is usually used to analyze it. Given that the dependent variable data contains lots of zeros, and the largest value is higher than a Poisson distribution predicts, the negative binomial regression is adopted.

3.3.2 The measurements of variables

Figure3-1 shows the distribution of patent applications in prefectures in China. The left map shows that innovation activity is agglomerated in the eastern areas of China in 2001. The right map shows that most of the innovation activities were sustaining agglomerated in the eastern area, but some prefectures in the middle area of China were growing the innovative capabilities in 2008. However, paying attention to the number of patent applications, it can be found that innovation rapidly grew in the metropolis such as Shenzhen, Beijing, Shanghai, where has the advantage of diversity.

Table 3-1 shows the definition of variables and their resource. The period of this chapter is 2001-2008, and it covers 286 prefectures of China. We try to explore what are the determinants of the innovative capability that brought about geographic distribution in this chapter. The summary statistics also are appended in the table 3-1.

Table 3-2 shows the correlation of variables. All the independent variables are positively correlated with the invention of patent applications. For checking the multicollinearity of independent variables, the variance inflation factor (VIF) test is adopted following Tavassoli and Carbonara (2013). The results show that all independent variables got a value between 0 and 1 in the test, indicating the multicollinearity does not matter the empirical results.

3.3.2.1 Dependent variable

There are many contributions in the literature concerning regional innovation capacity and make use of many data like R&D expenditure, the profit of new product or patent as proxies of innovation capacity. For the previous literature, measurements of innovation capacity are classified by the production process into innovative inputs, innovative outputs and innovative agents (Sirilli 1998, Godin 2002a, Godin 2002b, Ratanawaraha and Polenske 2007).

The mainly used innovative inputs data in the empirical research are R&D expenditure, employment in high-tech manufacturing, or creative sectors. For instance, Feldman and Lichtenberg (1998) using a database of funding for research projects in the EU to measure the spatial distribution of innovation. However, the OECD manual (2002, p.19) points out the error that possibly exists in measuring R&D is "the difficulty of locating the cut-off point between experimental development and the related activities required to realize innovation." This makes this measurement a limited quality when measuring innovation capacity.

The mainly used innovative outputs data in the empirical research are patent counts and new products counts. There are a number of previous works on innovation use patent counts as a proxy (Pakes et al. 1984, Acs et al. 2002, Buesa et al. 2010). Patents could reflect the technological change so that they are expected to indicate the new technology creation. However, Hall et al. (2001) argued that simple patent counts show the technological importance or value of inventions but don't show the economic importance or value of patents. To remedy this issue, patent citations are often used instead of simple patent counts. Hamaguchi and Kondo (2016) used Japanese patent citations as an indicator of high-quality innovation in their research to study the interregional knowledge turnover effects on the quality of innovation.

Albeit the innovation counts are used as proxies of innovation, it basically based on surveys of firms which make this kind of data is unavailable in some countries and exist a problem of time continuity. For example, Feldman (1994) used the number of new products introduced to the market to explore the spatial distribution of innovation activities in the U.S.

Innovative agents such as the R&D institutes, universities, new firms or publications on the R&D outcomes are used as proxies of innovation because the agents directly produce innovations (Adams and James 2002, Zucker et al. 1994, Feldman and Linchtenberg 1998). However, as Schumpeter considered, after implementing commercialization of the new product is called the innovation. The innovative agents are not a good proxy for innovation.

In conclusion, patent counts relatively have the advantage to be the proxy of innovation capacity. For the firms are much more motivated to apply patents (Buesa et al. 2010) due to the high cost in the process of applying for patents, patents are better to explain the innovation especially when studying in a regional level. Besides, a long time period of the patent database is available gives it an advantage in over other measurements. Thus, in this chapter, we use the patent data as the proxy of innovation capacity.

In this chapter, limited by the source of data, a kind of patent counts is adopted as a proxy of regional innovation capacity. The database of patent applications matched with ASIE is obtained from the Chinese Patent Data Project (He et al. 2018). In China, there are three kinds of patent include invention, external design, and utility model. Due to the limitation of data, it is hard to obtain information on the citation of patents. For controlling the quality of patents, in this chapter, only invention patents data are used to indicate the innovation capacity of prefectures. The database is a micro dataset with 332,682 records, in which the addresses of firms are contained. The number of invention patent applications in prefectural level are counted by year. In order to match with available data of regional information, data of minority autonomous prefectures are excluded.

3.3.2.2 Independent variables

As explained in the last section, The China City Statistical Yearbook only provides information about 286 prefectures, excluding minority autonomous prefectures. Thus, the number of prefectures in this chapter is 286. Based on the discussion in section 2 on the determinants of regional innovation capability, we employ the following measurement of these determinants.

(I) Variety of innovation: The Herfindahl-Hirschman Index (HHI) is often used to enumerate the degree of concentration of the market. In this chapter, the inverse of HHI is used to calculate the index of the variety of innovation. The patent database is classified into 2-digit industrial sectors and 286 prefectures by years. The index of the variety of innovation is given by:

$$VRT_{r,t} = 1/\sum_{i=1}^{N}(\frac{x_{ir,t}}{x_{r,t}})^2 \quad (10)$$

where $x_{ir,t}$ is the number of patent application in industrial sector i in region r at time t. And $x_{r,t}$ is the total number of patent applications in region r at time t. The higher value of HHI, the higher degree of concentration will be. In other words, the variety is lower when the value of HHI is high. For convenience, the inverse of HHI is used in the regression. The prefecture has more varieties of innovation is expected to be more innovative.

(II) Innovative inputs：According to the theory of knowledge production function, there is a correlation between innovative output and innovative input (Griliches 1979). In this study, two kinds of innovative inputs are considered. One is the government expenditures on R&D，the other is human capital.

The variable $RD_{r,t}$ in the model is calculated as the log of R&D expenditure in region r at time t. Following the literature, it is expected to have a positive effect on the innovative capability of prefectures.

The variable $HC_{r,t}$ is calculated as the number of college students in region r at time t. The region that has a higher proportion of educated students has a higher innovative capability.

(III) Market structure: The market structure has a powerful influence on the innovative capability of prefectures in the context of China. The market structure is measured in two ways: one is named marketization that is calculated by the ratio of the output of non-state enterprises to the total in region r at time t. The prefectures with higher marketization are more favorable to the innovative capability of prefectures. The other one is globalization that is calculated by the ratio of the output of foreign enterprises to the total in region r at time t. In the literature, FDI is shown that has a positive effect on the innovative capability of provinces. In the level of prefectures, it is expected to have a positive effect on the innovative capability as well.

(IV) NHIDZ dummy: To identify the difference of innovative capability affected by the NHIDZ, the NHIDZ dummy is used in the model. The function of NHIDZ to promote innovative capability is discussed in section 2. It can be expected that the prefecture that owns an NHIDZ has a higher innovative capability.

3.4 Data, empirical method and results

3.4.1 Data

The raw data of patent applications has 332,682 records. After dropping the records of firms located in the minority autonomous prefectures, we are left with 326,123 records to be used in this chapter. Using the reference table of international patent classification and national industrial classification published by SIPO, the 2-digit industrial classification is matched with the patent data. These records are also tagged by prefecture-level locaton indicator for each year. The dataset of this chapter is a panel data with 286 prefectures in the period of 2001-2008.

Due to the lack of information in the China City Statistical Yearbook to calculate the market structure, AISE database is used. Using both firms' address and ownership information included in AISE database, we can calculate the ratio of the output of different ownership to the total output.

3.4.2 Empirical results

We adopt the negative binomial regression to estimate equation (1). The results are reported in Table 3. The results of the random-effect model with a cluster-robust error are shown in column (1). The coefficients of R&D expenditure and human capital are positive and significant as expected. They show that innovative inputs have significant effects on the regional innovative capability. If a prefecture were to increase its log of R&D expenditure by 1%, the difference in the log of expected innovative capability would be expected to increase by 0.212unit, while holding the other variables in the model constant. In the same way, the increment of human capital gives a slight effect on the innovative capability by only 0.003 unit. The hypothesis 1 and 2 are confirmed.

The influence of marketization is positive and significant as expected, too. The result shows that the less local influence of the SOEs is in a prefecture, the stronger innovative capability of the prefecture will be. If a prefecture were to increase its marketization by 1%, the difference in the log of expected innovative capability would be expected to increase by 0.9 unit, while holding the other variables in the model constant. Hypothesis 3 is confirmed.

However, contrary to expectation, the estimated coefficient of globalization is negative and significant. That means hypothesis 4 is rejected. The result shows that if a prefecture were to increase its marketization by 1%, the difference in the log of expected innovative capability would be expected to decrease by 0.422 unit while holding the other variables in the model constant. It indicates that the expected technology diffusion from FDI is not relevant at the prefecture level.

Variety of innovation is positive and significant as expected. If a prefecture were to increase its variety of innovation by 1%, the difference in the log of expected innovative capability would be expected to increase by 0.106 unit, while holding the other variables in the model constant. Hypothesis 5 is confirmed.

NHIDZ is positive and significant as expected. If a prefecture were to set up an NHIDZ, the difference in the log of expected innovative capability would be expected to increase by 0.233 unit, while holding the other variables in the model constant. Hypothesis 6 cannot be rejected in the random-effect regression.

The intercept term is negative and significant. That means when all independent variables in the model are evaluated at zero, the log of the expected count for the innovative capability of prefectures is -2.193 unit.

The results of the fixed-effect model with a cluster-robust error are shown in column (2). The effects of the log of R&D and Human capital are similar to the results of the random-effect model. But other variables are different in the effectiveness. The effect of marketization is stronger than that estimated in the random-effect model.

Globalization is negative and significant. It is robust with the random-effect model but contrary to expectation. The result means that if a prefecture were to increase its globalization by 1%, the difference in the log of expected innovative capability would be expected to decrease by 0.942 unit while holding the other variables in the model constant. Hypothesis 4 is rejected.

Variety of innovation is positive and significant as same as in the RE model. But its effectiveness is smaller than that in the RE model. Only 0.083unit log of expected innovative capability would be expected to increase if the variety of innovation were to increase by 1%. And NHIDZ is insignificant in the FE model.

The result of the LR test with a pooled negative binomial model is reported in Table.4 as well. It shows that the null hypothesis of alpha equal to zero is clearly rejected. That means the panel negative binomial regression is a preferred choice. Besides, because of the dispersion parameter r and s are significantly over 0, it means that the overdispersion exists and is better to use a negative binomial model than a Poisson model (see the manual of STATA of command xtnbreg).

3.4.3 The results including the interaction term

In order to check the effect of NHIDZ on the innovative capability of prefectures better, an interaction term of NHIDZ and variety of innovation is included in the models and the results are shown in Table.4.

Column (1) shows the results of the RE model, including the interaction term. Comparing with the result excluding the interaction term, most variables exhibit their effectiveness on the innovative capability of prefectures. But the effectiveness of NHIDZ has a significant difference from the one excluding interaction term. In the RE model, the coefficient of NHIDZ changes from 0.233 to 1.092 while remaining statistically significant. In the FE model, the coefficient of NHIDZ changes from 0.04 to 0.793 and becomes significant.

The interaction term of NHIDZ and variety of innovation is negative and significant both in the RE model and the FE model. Checking the z-statistic on the interaction term, it is -3.35 in the RE model and -3.46 in the FE model. It indicates that this interaction term improves the goodness of fit of the model (Karaca-Mandic et al. 2012). When NHIDZ changes from 0 to 1, will increase the log of expected innovative capability, but the effectiveness will be decreased by the cross partial effect. At the same time, the effectiveness of the variety of innovation on the innovative capability of prefectures will decrease when a prefecture owns an NHIDZ.

3.4.4 A discussion of global spatial autocorrelation

The analysis of this chapter implies a premise that all regions have a spatial random distribution.

But as shown in Figure 3-1, it seems that the spatial autocorrelation exists in the concentrated regions. Thus, a test for global spatial autocorrelation needs to be held. The Moran's I is a widely used method to test that (Cliff and Ord 1970; Anselin 1995). In this chapter, the Moran's I of each year in the period are calculated and reported in Table 3-5. All results are insignificant. It suggests that the distribution of the innovation creativities cannot reject a spatial randomness hypothesis.

3.5 Conclusion

This chapter analyzed the effect of (i) NHIDZ, (ii) variety of innovation, (iii) innovative inputs and (iv) market construction on the innovative capability of prefectures in China. The results of the empirical analysis show that innovative inputs give a positive impact on the innovative capability of prefectures. The marketization affects the innovative capability of prefectures strongly in a positive way. However, the effect of globalization is contrary to the expectation that has a negative effect on the innovative capability. Variety of innovation has a positive effect on the innovative capability of prefectures and NHIDZ does the same. However, the interaction term of them is negative and significant, indicating the cross partial effect weaken the marginal effects of NHIDZ and variety of innovation when they influence the innovative capability at the same time.

The conclusion of this chapter could be summarized as follows:

(I) A prefecture with more innovative inputs such as R&D expenditure or human capital has a better expression on the innovation activities. Some inland prefectures with a relatively abundant innovative resources could be expected to raise the innovative capability in the long term.

(II) The market construction affects the innovative capability. The higher marketization of a prefecture, the high the innovative capability of the prefecture will be. The policy implication of this result is that police makers should promote marketization and reduce subsidies to SOEs. But the result of effect of foreign firms which is expected to have a knowledge spillover on the innovative capability of prefectures, showing a contrary result with the previous literature (Chueng and Lin 2004). Policymakers usually believe in that FDI could improve the local innovative capability so that many preference policies for attracting FDI are made, but the result in this chapter shows that more FDI output shares is give a negative effect on the innovative capability.

(III) Variety of innovation impact upon the innovative capability of prefectures positively. The diversity of innovation improves the prefectural creativity indicates that police maker should encourage all kinds of industries to be innovative, rather than supporting single industry.

(IV) NHIDZ has a positively effect on the innovative capability of prefectures as expected. These preference policies for the NHIDZ worked on the improvement of prefectural innovative capability. However, the endogenous problem disturbs this conclusion in this chapter. To deal with it, a method of instrumental variables will be in the next stage.

The interaction term of NHIDZ and variety of innovation show us the limitation of policy implication. The NHIDZ could impact upon the prefectural innovative capability effectively, but the effectiveness will be discount by the cross partial effect of NHIDZ and variety of innovation. NHIDZ has a problem which has been argued in many pieces of literature is the industrial convergence in most of NHIDZs in China. It may can explain why the interaction term is negative.

(V) This chapter explored the determinants of the innovative capability of prefectures. It highlights the effect of NHIDZ on the prefectural innovative capability and found that the NHIDZ really has a positive effect on increasing the innovative capability, but the effectiveness will be discount by the cross partial effect. It is speculated that the result is related to industrial convergence problem. Remind the policymakers to need to coordinate the relationship between industrial diversification and convergence when making policies to promote the regional innovative capability.

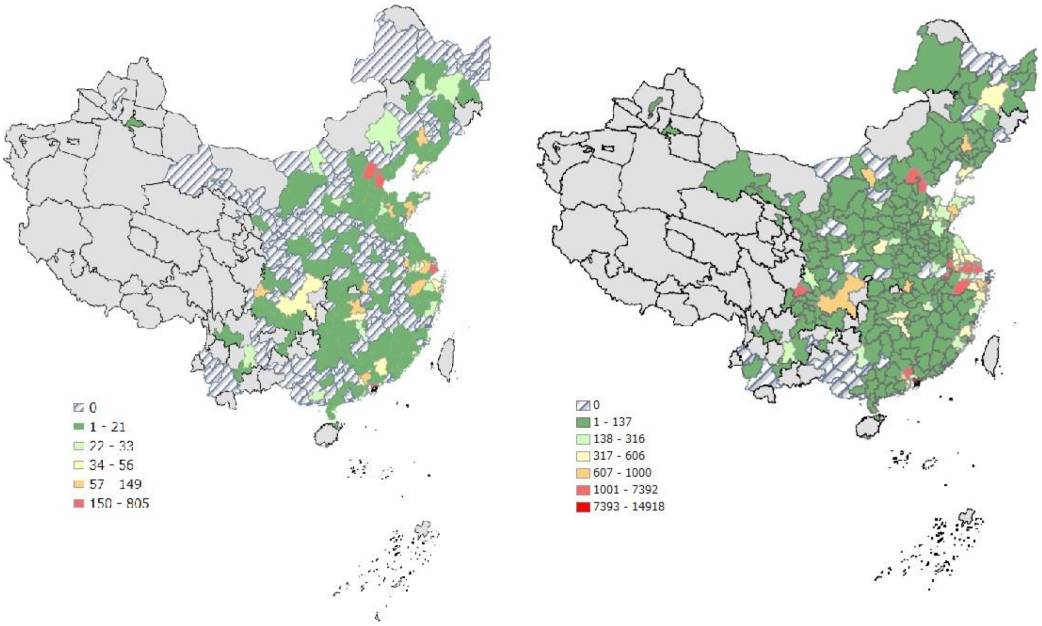


Figure 3-1. Distribution of the number of patent applications over Chinese prefectures (2001 and 2008)

Table3-1. Definition of variables and summary statistics

Variables	Definition	Source	unit	Mean	Std.	Min	Max
Innovation capacity	Number of prefectural granted patent	SIPO	number	138.831	756.384	0	14968
R&D	Prefectural fiscal expenditure on science	China City Statistic Yearbook	million yuan	6335.733	39761.480	0.130	1057666
Human Capital	The number of college students in prefectures	China City Statistic Yearbook	1,000 person	55.877	102.393	0	778.368
Marketization	The ratio of non-state enterprises output to the total output in prefectures	ASIE	%	0.790	0.196	0.024	1
Globalization	The ratio of foreign enterprises output to the total output in prefectures	ASIE	%	0.182	0.190	0	0.960
Industrial zone	Dummy of the National High-tech Industrial Development Zone	MOST	unit	0.247	0.431	0	1
Innovation variety	the inverse of HHI which is calculated by sectoral patent number in prefectures	SIPO		3.554	2.333	1	12.941

Table 3-2. Correlation matrix

Variables	PATENT	RD	HC	MKT	GLZ	NHIDZ	VRT
PATENT	1						
RD	0.573	1					
HC	0.292	0.3462	1				
MKT	0.105	0.0807	-0.0039	1			
GLZ	0.307	0.2086	0.2441	0.3806	1		
NHIDZ	0.148	0.1667	0.5892	-0.053	0.2764	1	
VRT	0.140	0.1882	0.4932	0.1566	0.3108	0.4884	1

Table3-3. Estimation results. Output : number of patent applications

	(1)	(2)
	RE neg-bin panel	FE neg-bin panel
RD	0.212***	0.212***
	(0.018)	(0.016)
HC	0.003***	0.003***
	(0.000)	(0.000)
MKT	0.900***	1.281***
	(0.184)	(0.171)
GLZ	-0.422*	-0.942***
	(0.237)	(0.266)
VRT	0.106***	0.083***
	(0.021)	(0.016)
NHIDZ	0.233*	0.04
	(0.134)	(0.139)
CONS	-2.193***	-2.18***
	(0.172)	(0.150)
ln r	0.290*	
	(0.149)	
ln s	2.373***	
	(0.248)	
LR test vs. pooled	2087.84	
	(0.000)	
Observations	1677	1644
Wald chi(2)	1070.79	894.85
	(0.000)	(0.000)
Log-likelihood	-7095.123	-5426.72

Cluster-robust errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table3-4. Estimation results. Output : number of patent applications (added interaction term)

	(1)	(2)
	RE neg-bin panel	FE neg-bin panel
RD	0.266***	0.255***
	(0.018)	(0.019)
HC	0.004***	0.004***
	(0.000)	(0.000)
MKT	0.832***	1.217***
	(0.193)	(0.166)
GLZ	-0.412*	-0.947***
	(0.238)	(0.273)
VRT	0.106***	0.083***
	(0.021)	(0.016)
NHIDZ	1.092***	0.793***
	(0.278)	(0.311)
VRT×NHIDZ	-0.121***	-0.107*
	(0.036)	(0.039)
CONS	-2.502***	-2.418***
	(0.173)	(0.176)
ln_r	0.307***	
	(0.148)	
ln_s	2.374***	
	(0.245)	
LR test vs. pooled	2076.370	
	(0.000)	
Observations	1677	1644
Wald chi(2)	1210.540	960.230
	(0.000)	(0.000)
Log-likelihood	-7084.063	-5418.852

Cluster-robust errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table3-5. Moran's I statistic for invention patent applications of prefectures

	Moran's I	E(I)	SE(I)	Z(I)	p-value
2008	0.040	-0.004	0.019	2.340	0.019
2007	0.021	-0.004	0.016	1.518	0.129
2006	0.017	-0.004	0.016	1.300	0.194
2005	0.019	-0.004	0.018	1.270	0.204
2004	0.021	-0.004	0.021	1.180	0.238
2003	0.019	-0.004	0.021	1.078	0.281
2002	0.028	-0.004	0.022	1.453	0.146
2001	0.031	-0.004	0.021	1.600	0.110

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

Chapter 4 Localization of Innovative Activities in Yangtze River Delta, China

4.1 Introduction

Previous studies show that are innovative activity more concentrated in some points than the general economic activities (Carlino and Kerr, 2014; Buzard and Carlino, 2013; Chatterji et al., 2014; Fornahl and Brener, 2009). As a means of gaining competitiveness, innovation capability is important for both firms and economic development of regions. The issues of innovation received many attentions of scholars. Some studies focus on the internal organization structure, the R&D effort and other factors affecting on the innovation capability of firms (Calantone et al., 2002; Romijn and Abaladejo, 2002). Other studies pay attention to the regional innovation capability following Marshall's (1890) argument of a knowledge spillover. In this view, location matters to firms' innovative activity (Asheim and Gertler, 2006; Lai et al., 2015; Buesa et al., 2010).

Due to the limitation of data and other conditions, previous research on innovative activity have been carried out either at macro-level or specific case studies. Recently, owing to wider availability micro-level data makes it possible to grasp the innovative activity at a business unit level (Fornahl and Brener, 2009). However, the innovative activity in the developing countries are still unclear. Most empirical researches on developing countries are done at a macro level. The emerging countries like China shows not only a rapid economic growth but also significant technology development. It is important to investigate the factors affecting innovative activity of China's manufacturing industries by using a micro data.

To fill this gap, this chapter aims to explore the localization of innovative activity of industries in

China. Although China is a big country, the innovative activity tends to usually take place in a narrow space. Our study focuses on the Yangtze River Delta (YRD) region. Yangtze River Delta region is one of China's most developed regions with intense industrial agglomeration. Also, the Yangtze River Delta is the most active region in innovative activity. According to the Patent Statistics Brief (2018) published by SIPO, the spatial distribution of the number of patent activities, including the number of patent applications, authorizations, valid patents and PCT patent applications, are concentrated in the Yangtze River Delta and the Pearl River Delta. Therefore, we can expect that analysis of the localization of innovative activity in Yangtze River Delta will be very insightful to understand the nature of innovative activity of China.

We use a Chinese patent database available in by Harvard Dataverse to measure the innovation. For a detailed analysis of the localization of innovative activity in YRD region, we conduct sector-by-sector analysis.

The distribution of China's innovative activity has been examined by some empirical studies. Most studies use a provincial data and found that the innovative activity in China are unevenly distributed in the eastern developed provinces. (Sun, 2000, 2003; Cao and Qin, 2012; Bickenbach and Liu, 2014) Recently, Ma and Liu(2019) studied the concentration of innovative activity in China focusing on the three mega-economic zones of Beijing-Tianjin-Hebei(JJJ), Yangtze River Delta(YRD) and Pearl River Delta(PRD) by using several types of innovation proxies like patent applicants, R&D expenditure and other variables.

Studies on the distribution of innovative activity in developed countries find that the innovative activity are more concentrated in some specific areas than general manufacturing. We will ask whether the same conclusion can be obtained in Chinese case. In the existing literature, the concentration of innovative activity is measured by various scalar indices such as local Gini coeffecient, EG index and General Theil index. However, these methods are fundamentally aspatial without any expression of spatial dimension; therefore, it is hard to get the image of actual location of innovative activity.

To examine the concentrated degree of innovative activity and to get a spatial image of the localization of innovative activity in YRD, following Duranton and Overman (2005), this chapter uses a spatial continuous-distance approach to explore the detailed localization of industrial innovation in YRD. This approach is used in this chapter basing on the idea of comparing the distribution of distances between pair of innovation enterprises in an industry to the hypothesis that their locations are randomly chosen from the locations of all enterprises in the same industry in YRD.

The results show that the tendency of innovative activity to localization is not appeared in all industries that enterprises are actively innovating. And some findings are interesting. For instance, the cotton textile and printing and dyeing finishing is regarded as a labor-intensive industry but its innovative activity are concentrated than general economic activities. Also, we found that the localization of innovative activity is independent of its localization of general economic activities in the same industry.

The rest of the paper is organized as follows. In the next section we will introduce the database we use and how we construct our data set. The methodology of our study will be presented in section 3. Section 4 presents the results and we conclude in section 5.

4.2 Data

4.2.1 Data source

The Yangtze River Delta region includes 11 prefecture-level cities in Shanghai and Zhejiang Province and 13 prefecture-level cities in Jiangsu Province. Our analysis uses the manufacturing enterprises in the Yangtze River Delta region. The manufacturing industry consists of the two digits industry classification of 13-37 and 38-43 in the National Economic Industry Classification (GB/T 4754-2002).

In this chapter we use a micro data set from two database: the patent database from the State Intellectual Property Office (SIPO) and the Annual Survey of Industrial Enterprises (ASIE) that administered by the National Bureau of Statistics of China. This database is provided by He et al. (2017) at the Harvard Dataverse. Because their patent database only contains the information on patents applied by enterprises and the industrial enterprises number. In order to obtain more information of enterprises we matched the patent database with ASIE by ourselves.

Since ASIE database only collected the data of state-owned enterprises and collective-owned enterprises before 1998 and the criteria of inclusion for firms changed in 2011, the period of the matched patent database is restricted in 1998-2009 in order to cover the ASIE data. See He et al. (2017) for a detailed explanation.

Considering innovative activity is an ongoing process, this chapter pools the 1998-2009 manufacturing companies in the Yangtze River Delta region and removes duplicate values based on the company's code and name. After matching the last two databases, a total of 176,251 enterprise data was obtained, including 8585 patent-pending companies. Also, considering that the Yangtze River

Delta region is one of the most economically intensive regions in China, it is necessary to select more detailed classification in order to better grasp the agglomeration of innovative activity in different industries. However, due to the difference in innovative activity in different industries, too detailed classification will cause some industries without data on innovative activity. Taking into account the above situation, this chapter divides the manufacturing industry into three-digit classification.

For the confidence of estimated result, we selected these industries with sufficient level of innovative activity to estimate their kernel density function. We selected industries having more than 1% of the number of patent-applicants of ll firms in 1998-2009. Table 1 presents a list of those industries. From the list we can see that there are 34 industries are active in innovative activity and most industries belong to technology-intensive industry or capital-intensive industries but some labor-intensive industries appear in the list, as well.

4.2.2 Space coordinate conversion processing

Our analysis requires the distance between firms as an important variable, we obtain geographic coordinates from the enterprises' address information. We use XGeocoding software for geocoding and the Gaode map to query coordinates information.

Figures 4-1(a-d) are map of location information for four industries which innovative activity are localized and have the most active innovative activity in YRD: Manufacture of Special Chemical Products(SIC266), Manufacture of pumps, valves, compressors and similar machinery(SIC354), Manufacture of Vehicles(SIC372), Manufacture of Electronic industrial apparatus(SIC392). The dots present the location of innovation enterprise and crosses present the location of other enterprises without patent application in this period in the same industry. Each map shows that the innovative activity is more concentrated relative to the distribution of general economic activities in the same industry. It also can be seen that most innovative activity are localized at Shanghai or cities near to Shanghai. In next section, the methodology we use to evaluate the degree of localization will be introduced. It could make us a more precise image of the concentration of industrial innovative activity.

4.3. Methodology

In this chapter, we refer to the spatial kernel density estimation method proposed by Duranton and Overman (2005) and use R language to program the calculation process and then simulate the result.

4.3.1 Estimating K-density

In this chapter, to assess the concentration of innovation enterprises with regard to their industry, the estimation process is subdivided into the following steps: Firstly, using the latitude and longitude data of each enterprise prepared in advance, the distance matrix after pairing between each two enterprises is calculated. The number of enterprises with innovative activity of a specific industry is n, so it has a total of $n*(n-1)/2$ different distance pairs. The kernel density function of the bilateral distances at any distance d is given by equation (11):

$$\hat{k}(d) = \frac{1}{n(n-1)h}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} f\left(\frac{d-d_{ij}}{h}\right) \quad (11)$$

where d_{ij} is the distance between enterprise i and j, h is the bandwidth which is calculated following Silverman (1986).[3] f is a Gaussian kernel function. This function will be simulated relying on randomly extracting 512 distance-pairs by Monte-Carlo simulations 1,000 times.

4.3.2 Counterfactual and confidence band

The second step is to construct a counterfactual and define the confidence bands to compare the actual k-densities to the counterfactuals. In this chapter, the counterfactual is constructed as the location of innovation enterprises to be randomly chosen from all the location in the same industry including innovation enterprises and general enterprises. It helps us to look at the localization degree of innovative activity in an industry relative to the localization degree of that industry. Under the hypothesis, the k-density function is estimated 1,000 times to obtain the points that are included in the 95% confidence intervals in the same way we did for the actual locations.

Following Duranton and Overman (2008), we firstly simulate the k-densities over all possible distances. Then comparing the actual k-density with the counterfactual k-densities, if for short distances there are 'abnormally' high values for the distance density which is bigger the counterfactual density then the $\hat{k}(d)$ could be interpreted as localization. On the contrary, if smaller than the counterfactual density in the short distances then the $\hat{k}(d)$ could be interpreted as dispersion.

Now we present specific definition and calculation methods for confidence band. Following Duranton and Overman (2005, 2008), we define and compute both local confidence bands and global confidence bands. We define an upper 2.5% and a lower 2.5% confidence level for distance d, it makes that 2.5% of the randomly generated k-densities lie above it at distance d, similarly 2.5% of

---

[3] In this chapter, the bandwidth is calculated as: $h = 0.9 * \left(\frac{dis_{0.75} - dis_{0.25}}{1.34}\right) * n^{-\frac{1}{5}}$, *in which*

which lie below it at distance d. At this point, we have local confidence bands, but this only allow us to observe the situation for a given distance d. Following Duranton and Overman (2005, 2008), we only pay attention to the short distances which the threshold is decided by the median of all distances that between two innovation firms over all industries. Through calculating we got a threshold of 158 kilometers, so we draw global confidence bands as the sum of distances in the interval between 0 and 158 kilometers.

To compute global confidence bands, we denote the upper global confidence band by $\bar{k}(d)$ for innovation enterprises in an industry. If there exists at least one $d \in [0,158]$ makes $\hat{k}(d) > \bar{k}(d)$ established, the innovation enterprises in this industry could be said to exhibit localization (at a 5% confidence level). Similarly, we denote the lower confidence band by $\underline{k}(d)$. If $\underline{k}(d) < \hat{k}(d)$ for at least one $d \in [0,158]$ and they don't exhibit localization then we could say that the innovation enterprises in this industry exhibits no stronger localization vis-à-vis general tendency of same industry. Based on the definition, we could compute the degree of localization for the innovation enterprises of an industry. As this method is presented by Duranton and Overman (2005), we call the index by DO index. The indexes of localization and dispersion are defined as follows:

$$\Gamma(d) \equiv \max\left(\hat{k}(d) - \bar{k}(d), 0\right), \tag{12}$$

$$\Psi(d) \equiv \begin{cases} \max\left(\underline{k}(d) - \hat{k}(d), 0\right) if \ \sum_{d=0}^{d=158} \Gamma(d), \\ 0 \qquad\qquad\qquad\qquad\qquad otherwise. \end{cases} \tag{13}$$

The localization of innovative activity in an industry is detected when the k-density lies above the upper global confidence band. For the situations the k-density lies below the lower global confidence band or never above the upper global confidence band, it is detected as dispersion of innovative activity in an industry.

4.4 Results

The results of 34 industries that are the most active in innovative activity present in Table 4-2. From Table 4-2 we can see that not in all industries the innovation enterprises are more concentrated relative to the general enterprises in the same industry. Among the selected 34 industries, 15 industries their innovative activity are localized in a short distance. However, there are 19 industries their innovative activities are active but no stronger localizaiton.

When we look at the results of industries their innovative activities are localized, we found that there are two patterns of localization in these cases. The one is innovative activity are only concentrated at a certain range of distances in the interval [0,158]. Printing, textile processing and Dyeing of Cotton & Chemical Fibers (SIC171), Textile Products Manufacturing (SIC175), Synthetic Material Manufacturing (SIC265), Special Chemical Product Manufacturing (SIC266), Glass and Glass Products Manufacturing (SIC314), Non-ferrous Metal Rolling (SIC335), Structural Metal Products Manufacturing (SIC341), Fan, Weighing and Packaging Equipment Manufacturing (SIC357), General Parts Manufacturing and Mechanical Repair (SIC358), Electrical Manufacturing (SIC391), Transmission and Distribution and Control Equipment Manufacturing (SIC392), Manufacture of Wires, Cables, Cables and Electrical Equipment (SIC393), Communication Equipment Manufacturing (SIC401) have a tendency of cluster in a certain range of distances. Among these, Printing, textile processing and Dyeing of Cotton & Chemical Fibers (SIC171), Glass and Glass Products Manufacturing (SIC314), Structural Metal Products Manufacturing (SIC341) have a relatively far distance in concentration which are more than 100km.

The other one is innovative activity are concentrated in two ranges of distances in the interval

[0,158]. Vehicle Manufacturing (SIC372), Lighting Equipment Manufacturing (SIC397) have a tendency of cluster in two certain range of distances.

To show the localization of innovative activity in an industry graphically, we illustrate the same four industries are refered in Section 2 in Figure 2 (a-d). The actual k-density is presented by the blue line, local confidence bands by the black dotted lines and global confidence bands by black lines. The results from those figures consist with information conveyed on the maps in Figure 4-2(a-d). Innovative activity in Special Chemical Product Manufacturing (SIC266), Vehicle Manufacturing (SIC372), Transmission and Distribution and Control Equipment Manufacturing (SIC392) are localized relative to the industries. However, even though Pumps, Valves, Compressors and Similar Machinery Manufacturing (SIC354) is active in innovation, its innovative activity is dispersed in YRD.

In addition, due to YRD is one of economic activities are the most agglomerated areas in China, in 3 digits classification most industries are localized, it is curious whether the agglomeration of economic activities is related to the concentration of innovative activity. In order to examine this, we compute the DO index of the 16 industries that their innovative activities are localized with the counterfactual enterprises in these industries are randomly chosen their location from all industries. By using the Spearman-rank correlation, we examine the correlation between the index of localization of innovative activity in the industry and that of localization for the industry relative to the whole industries (as Duranton and Overman did, 2005, 2008) and the result shows that it is insignificant. Thus, it can be concluded that the localization of innovative activity in an industry is not related to the localization of the industry to the whole.

The results show that not all industries with sufficient level of innovative activity tend to concentrate to take advantage of the externality of knowledge spillover. And the competitiveness of firms in the same industries seems have no impact on the localization of innovative activity of that industry.

4.5 Conclusion

In this chapter we explored the localization or dispersion of industrial innovative activity in YRD. This helps us to know more detailed facts of the localization of innovative activity in different industries. We find that not all industries that actively engage in innovative activity are localized. We also find in these localized industries, the concentration degrees are different and some even has two ranges exist concentration.

We also find an interesting result that Printing, textile processing and Dyeing of Cotton & Chemical Fibers (SIC171) which is regarded as a labor-intensive industry but its innovative activity is more localized relative to general production activities in the same industry. One possible explanation is the tacit knowledge plays an important role especially in the labor-intensive industry.

Then we check the correlation between the localization of innovative activity in an industry and that of localization for the industry relative to overall manufacturing and find that they are uncorrelated. This suggests that industrial agglomeration in China is not a factor in promoting the concentration of innovative activity.

As knowns, China's government has a great influence on industrial innovation through taxation, subsidies and other policies. It is interesting to look at if the localization or dispersion of innovative activity in an industry is affected by these policies. In future research, we will analyze the factors that influence the concentration of industrial innovative activity.

Table 4-1. The most active industries for innovation activities

SIC	Industry	Number	Proportion
171	Printing, textile processing and Dyeing of Cotton & Chemical Fibers	146	1.701
175	Textile Products Manufacturing	121	1.409
261	Basic Chemical Raw Materials Manufacturing	220	2.563
263	Pesticide Manufacturing	98	1.142
264	Paints, Inks, Pigments and Similar Products Manufacturing	202	2.353
265	Synthetic Material Manufacturing	136	1.584
266	Special Chemical Product Manufacturing	310	3.611
271	Chemical Drug Manufacturing	148	1.723
272	Chemico-pharmaceutial Preparations Manufacturing	123	1.433
314	Glass and Glass Products Manufacturing	107	1.246
335	Non-ferrous Metal Rolling	102	1.188
341	Structural Metal Poducts Manufacturing	103	1.200
352	Metal Processing Machinery Manufacturing	148	1.724
353	Hoisting Transportation Equipments Manufacturing	93	1.083
354	Pumps, Valves, Compressons and Similar Machinery Manufacturing	289	3.366
355	Bearings, Gears, Drives and Drive Components Manufacturing	98	1.142
357	Fan, Weighing and Packaging Equipment Manufacturing	230	2.679
358	General Parts Manufacturing and Mechanical Repair	110	1.281
361	Mine, Metalhurgy, Construction Equipment Manufacturing	94	1.095
362	Manufacturing of Special Equipment for Chemical, Wood and Non-metal Processing	188	2.190
365	Special Equipment Manufacturing for the Textile, Clothing and Leather Industries	133	1.549
368	Medical Equipment Manufacturing	94	1.095
369	Enviromental Protection, Social Public Safety and Other Special Equipment Manufacturing	132	1.538
372	Vehicle Manufacturing	309	3.599
391	Electrical Manufacturing	133	1.549
392	Transmission and Distribution and Control Equipment Manufacturing	340	3.960
393	Manufacture of Wires, Cables, Cables and Electrical Equipment	204	2.376
395	Household Electrical Appliance Manufacturing	122	1.421
397	Lighting Equipment Manufacturing	93	1.083
401	Communication Equipment Manufacturing	96	1.118
405	Electronic Device Manufacturing	167	1.945
406	Electronic Component Manufacturing	240	2.796
411	General Instrument Manufacturing	211	2.458

Note : The total number of innovation enterprises is 8585.

Table 4-2. Localized three-digit industries

SIC	Industiy	DO index	Localization range(km)
171	Printing, textile processing and Dyeing of Cotton & Chemical Fibers	0.012	[74,110]
175	Textile Products Manufacturing	0.031	[9,70]
265	Synthetic Material Manufacturing	0.005	[0,33]
266	Special Chemical Product Manufacturing	0.003	[0,25]
314	Glass and Glass Products Manufacturing	0.05	[0,125]
335	Non-ferrous Metal Rolling	0.003	[0,75]
341	Structural Metal Poducts Manufacturing	0.017	[115,157]
357	Fan, Weighing and Packaging Equipment Manufacturing	0.035	[0,69]
358	General Parts Manufacturing and Mechanical Repair	0.002	[0,25]
372	Vehicle Manufacturing	0.015	[7,61], [152,156]
391	Electrical Manufacturing	0.003	[45,54]
392	Transmission and Distribution and Control Equipment Manufacturing	0.001	[26,49]
393	Manufacture of Wires, Cables, Cables and Electrical Equipment	0.009	[0,31]
397	Lighting Equipment Manufacturing	0.064	[0,65], [88,124]
401	Communication Equipment Manufacturing	0.047	[0,40]

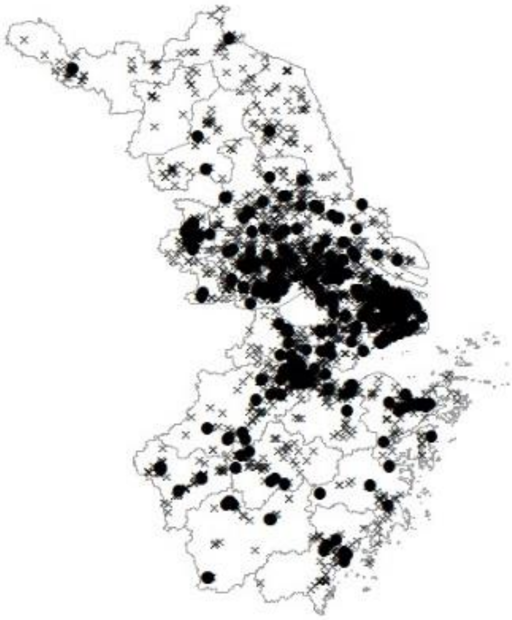


Figure 4-1. a. Printing, textile processing and Dyeing of Cotton & Chemical Fibers (SIC171)

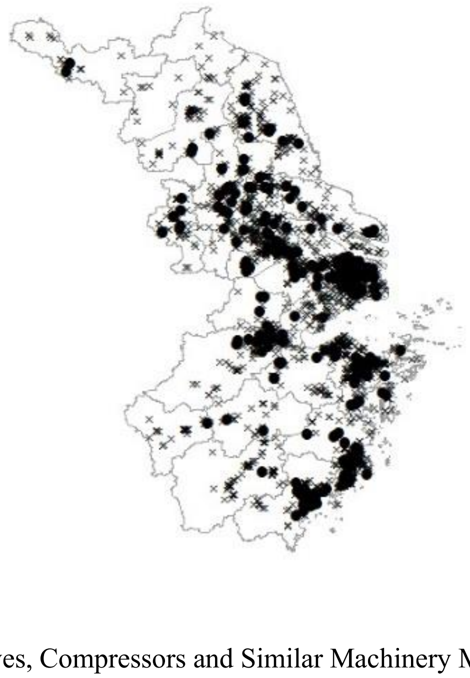


Figure4-1. b. Pumps, Valves, Compressors and Similar Machinery Manufacturing (SIC354)

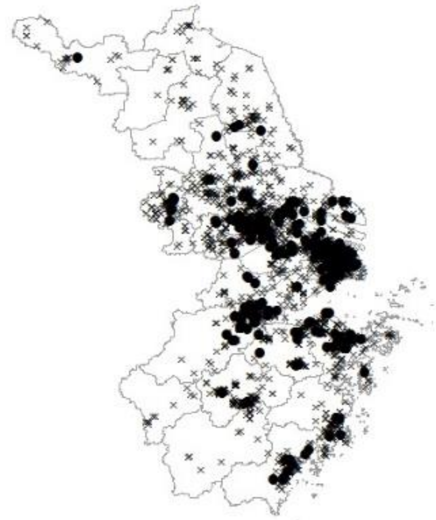


Figure4-1. c. Vehicle Manufacturing（SIC372）

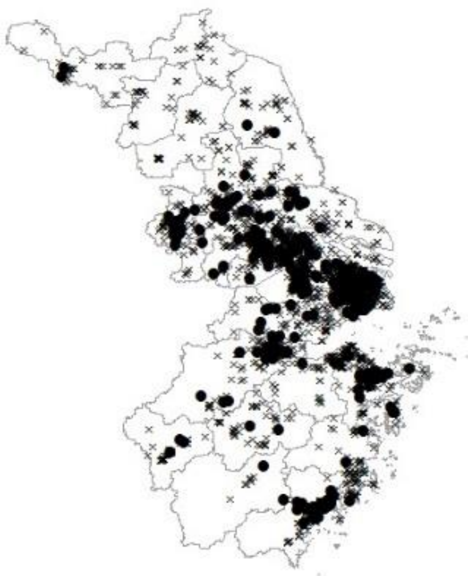
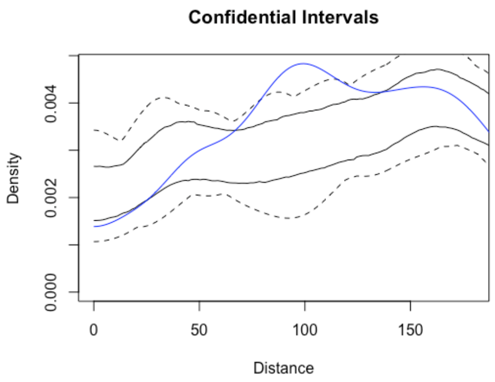
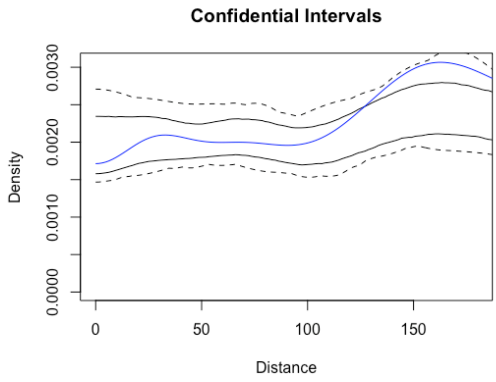


Figure4-1. d. Transmission and Distribution and Control Equipment Manufacturing（SIC392）

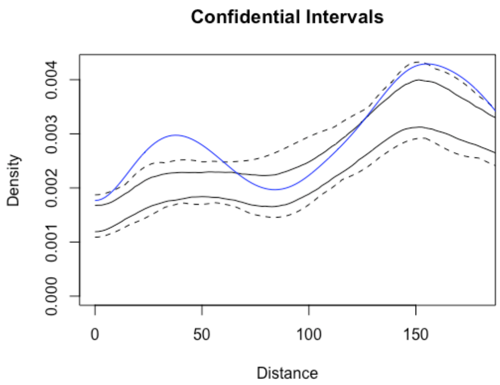
Figure 4-1. Maps of four illustrative industries



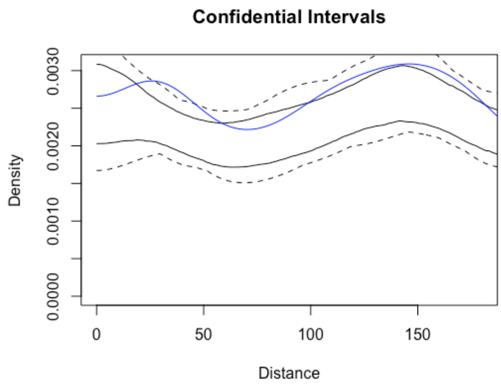
a. Printing, textile processing and Dyeing of Cotton & Chemical Fibers (SIC171)



b. Pumps, Valves, Compressors and Similar Machinery Manufacturing (SIC354)



c. Vehicle Manufacturing（SIC372）



d. Transmission and Distribution and Control Equipment Manufacturing（SIC392）

Figure 4-2. k-density, local confidence intervals and global confidence bands for four illustrative industries

Chapter 5 Conclusion

This study has investigated the importance of external environment and the characteristics of enterprises or industries to regional economic development. Our conclusion can be summarized as follows. Firstly, the role of externalities is fully demonstrated in our study. In Chapter 2, by exploring the factors affect the location choice of new firms in the Yangtze River Delta we find that firms make different choices under different ownership. Regional market potential is important to private firms but not to FDI. However, the congestion effect likes air pollution has a negative effect on both of them. In Chapter 3, we have examined that the diversity of knowledge plays an important role in affecting regional innovative capability. We also find that the construction of industrial zone may do not promote the regional innovative capability in the absence of other conditions. It indicates that external environment is essential for the development of regions.

Secondly, we can observe that different firms or industries react differently to the external environment so it is need to implement different policies for different objects. In Chapter 2, we considered the characteristics of firms then we find that the industrial agglomeration is important to FDI but not to private firms. In Chapter 4, we originally estimated the localization of innovation activities in different industries. The results show that not all industries have a tendency to concentrate the innovative activity to benefit from the knowledge spillover. The localized innovative activity of an industry is not related to the concentration of economic activity in the same industry in Yangtze River Delta.

Comparing with previous researches, our study has some improvements as follows: firstly, we use a micro data which provides more detailed information including spatial information and others. This allows us to explore issues from richer perspectives. Secondly, with the development of technology, some methods that were difficult to implement in the past, such as geocoding, can be applied to research easily now. This allows us to do more spatial exploration to better understand the regional economy. Lastly, the phenomena we have found provide a direction for further research.

In the future, we will develop a more detailed study of the factors affecting the localization of innovative activity in different industries.

Reference

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research policy, 31*(7), 1069-1085.

Adams, J. D. (2002). Comparative localization of academic and industrial spillovers. *Journal of Economic Geography*, *2*(3), 253-278.

Ağca, Ş., & Mozumdar, A. (2008). The impact of capital market imperfections on investment–cash flow sensitivity. *Journal of Banking & Finance*, *32*(2), 207-216.

Albahari, A., Catalano, G., & Landoni, P. (2013). Evaluation of national science park systems: A theoretical framework and its application to the Italian and Spanish systems. *Technology Analysis & Strategic Management*, *25*(5), 599-614.

Amable, B., & Petit, P. (2001). The Diversity of Social Systems of Innovation and Production during the 1990s N 2001-15. *Paris: Cepremap*.

Amiti, M., & Javorcik, B. S. (2008). Trade costs and location of foreign firms in China. *Journal of development economics*, *85*(1-2), 129-149.

Anselin, L. (1995). Local indicators of spatial association—LISA. Geographical analysis, 27(2), 93-115.

Arauzo‐Carod, J. M., Liviano‐Solis, D., & Manjón‐Antolín, M. (2010). Empirical studies in industrial location: an assessment of their methods and results. *Journal of Regional Science*, *50*(3), 685-711.

Audretsch, B. (1998). Agglomeration and the location of innovative activity. Oxford review of economic policy, 14(2), 18-29.

Audretsch, D. B., & Vivarelli, M. (1996). Firms size and R&D spillovers: Evidence from Italy. *Small Business Economics, 8*(3), 249-258.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review*, *86*(3), 630-640.

Audretsch, D. B., & Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In *Handbook of regional and urban economics* (Vol. 4, pp. 2713-2739). Elsevier.

Autant-Bernard, C. (2006). Where do firms choose to locate their R&D? A spatial conditional logit analysis on French data. *European planning studies*, *14*(9), 1187-1208. Carod, J. M. A., & Antolín, M. C. M. (2004). Firm size and geographical aggregation: an empirical appraisal in industrial location. *Small Business Economics*, *22*(3-4), 299-312.

Berliant, M., & Fujita, M. (2011). The dynamics of knowledge diversity and economic growth. *Southern Economic Journal*, *77*(4), 856-884.

Berliant, M., & Fujita, M. (2008). Knowledge creation as a square dance on the Hilbert cube. International Economic Review, 49(4), 1251-1295.

Bickenbach, F. and W. H. Liu. (2014), The development of the spatial concentration and co-agglomeration of industrial and public innovation in China – A province-level analysis. In I. Liefner and Y. D. Wei, eds, *Innovation and Regional Development in China*, New York: Routledge, 121–44.

Buesa, M., Heijs, J., & Baumert, T. (2010). The determinants of regional innovation in Europe: A combined factorial and regression knowledge production function approach. *Research policy*, *39*(6), 722-735.

Buzard, K., & Carlino, G. (2013). 16 The geography of research and development activity in the US. *Handbook of Industry Studies and Economic Geography*, 389.

Calantone, R. J., Cavusgil, S. T., & Zhao, Y. (2002). Learning orientation, firm innovation capability, and firm performance. *Industrial marketing management*, *31*(6), 515-524.

Carrincazeaux, C., Lung, Y., & Rallet, A. (2001). Proximity and localisation of corporate R&D activities. *Research Policy*, *30*(5), 777-789.

Cao, Y. and Y. X. Qin. (2012), Measurable analysis on the difference changes of regional innovation capability based on the dynamic performance in China. *Zhongguo Renkou Ziyuan Yu Huanjing* (*China Population, Resources and Environment*), 22(3), 164-9.

Carlton, D. W. (1979). *Why new firms locate where they do: An econometric model* (pp. pp-13). Cambridge, MA: Joint Center for Urban Studies of MIT and Harvard University.

Carlino, G., & Kerr, W. R. (2014). *Agglomeration and innovation* (No. w20367). National Bureau of Economic Research.

Carree, M. A., & Thurik, A. R. (1999). *Industrial structure and economic growth* (pp. 86-110). Cambridge, UK: Cambridge University Press.

Chatterji, A., Glaeser, E., & Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innovation Policy and the Economy*, *14*(1), 129-166.

Chang, S. J., & Park, S. (2005). Types of firms generating network externalities and MNCs' co‐location decisions. *Strategic Management Journal*, *26*(7), 595-615.

Chen, Y. (2009). Agglomeration and location of foreign direct investment: The case of China. *China economic review*, *20*(3), 549-557.

Cheung, K. Y., & Ping, L. (2004). Spillover effects of FDI on innovation in China: Evidence from the provincial data. *China economic review*, *15*(1), 25-44.

Cheng, S., & Stough, R. R. (2006). Location decisions of Japanese new manufacturing plants in China: a discrete-choice analysis. *The Annals of Regional Science*, *40*(2), 369-387.

Cheng, L. K., & Kwan, Y. K. (2000). What are the determinants of the location of foreign direct investment? The Chinese experience. *Journal of international economics*, *51*(2), 379-400.

Cheng, S., & Stough, R. R. (2006). Location decisions of Japanese new manufacturing plants in China: a discrete-choice analysis. *The annals of regional science*, *40*(2), 369-387.

Cliff, A. D., & Ord, K. (1970). Spatial autocorrelation: a review of existing and new measures with applications. Economic Geography, 46(sup1), 269-292.

Cohen, W. M., & Levin, R. C. (1989). Empirical studies of innovation and market structure. *Handbook of industrial organization*, *2*, 1059-1107.

Crozet, M., Mayer, T., & Mucchielli, J. L. (2004). How do firms agglomerate? A study of FDI in France. *Regional Science and Urban Economics*, *34*(1), 27-54.

Crescenzi, R., & Rodríguez-Pose, A. (2011). *Innovation and regional growth in the European Union*. Springer Science & Business Media.

Disdier, A. C., & Mayer, T. (2004). How different is Eastern Europe? Structure and determinants of location choices by French firms in Eastern and Western Europe. *Journal of comparative Economics*, *32*(2), 280-296.

Du, J., Lu, Y., & Tao, Z. (2008). Economic institutions and FDI location choice: Evidence from US multinationals in China. *Journal of comparative Economics*, *36*(3), 412-429.

Du, J., Lu, Y., & Tao, Z. (2012). Institutions and FDI location choice: The role of cultural distances. *Journal of Asian Economics*, *23*(3), 210-223.

Duoduo Tan, Cheng Cheng, Mujun Lei & Yucheng Zhao (2017). Spatial Distributions and Determinants of Regional Innovation in China: Evidence from Chinese Metropolitan Data. Emerging Markets Finance and Trade,53:6, 1442-1454.

Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *The Review of Economic Studies*, *72*(4), 1077-1106.

Duranton, G., & Overman, H. G. (2008). Exploring the detailed location patterns of UK manufacturing industries using microgeographic data. Journal of Regional Science, 48(1), 213-243.

Duranton, G., & Puga, D. (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review*, *91*(5), 1454-1477.

Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in US manufacturing industries: a dartboard approach. *Journal of political economy*, *105*(5), 889-927.

Fan, P. (2011). Innovation capacity and economic development: China and India. Economic change and restructuring, 44(1-2), 49-73.

Feldman, M. P. (1994). *The geography of innovation* (Vol. 2). Springer Science & Business Media.

Feldman, M. P., & Audretsch, D. B. (1999). Innovation in cities: Science-based diversity, specialization and localized competition. *European economic review*, *43*(2), 409-429.

Feldman, M. P. (1994). Knowledge complementarity and innovation. *Small business economics*, *6*(5), 363-372.

Feldman, M., & Lichtenberg, F. R. (1998). The interaction between public and private R&D investment: cross-country evidence from European Community's R&D information service. *Annales d'Economie et de Statistique*, *49*, 199-222.

Felsenstein, D. (1994). University-related science parks 'seedbeds' or 'enclaves' of innovation? *Technovation*, *14*(2), 93-110.

Fornahl, D., & Brenner, T. (2009). Geographic concentration of innovative activity in Germany. *Structural Change and Economic Dynamics*, *20*(3), 163-182.

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional studies*, *41*(5), 685-697.

Friedman, J., Gerlowski, D. A., & Silberman, J. (1992). What attracts foreign multinational corporations? Evidence from branch plant location in the United States. *Journal of Regional science*, *32*(4), 403-418.

Fujita, M., Krugman, P. R., & Venables, A. J. (2001). *The spatial economy: Cities, regions, and international trade*. MIT press.

Fujita, M., Hamaguchi, N., Kameyama, Y. (2018). *Spatial Economics in the Age of Declining Population.* Nikkeibook.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of* *political economy*, *100*(6), 1126-1152.

Godin, B. (2002). Outline for a history of science measurement. *Science, technology, & human values*, *27*(1), 3-27.

Godin, B. (2002). The rise of innovation surveys: Measuring a fuzzy concept. *Canadian Science and Innovation Indicators Consortium, Project on the History and Sociology of S&T Statistics, Paper*, *16*.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The bell journal of economics*, 92-116.

Hall, B. H. (1992). *Investment and research and development at the firm level: does the source of financing matter?* (No. w4096). National bureau of economic research. Himmelberg, C. P., & Petersen, B. C. (1994). R & D and internal finance: A panel study of small firms in high-tech industries. *The Review of Economics and Statistics*, 38-51.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: Lessons, insights and methodological tools* (No. w8498). National Bureau of Economic Research.

Hamaguchi, N., & Kondo, K. (2016, March). Fresh brain power and quality of innovation in cities: Evidence from Japanese patent database. In *Workshop on "Geography*.

Hansen, E. R. (1987). Industrial location choice in Sao Paulo, Brazil: a nested logit model. *Regional science and Urban economics*, *17*(1), 89-108.

Harris, C. D. (1954). The, Market as a Factor in the Localization of Industry in the United States. *Annals of the association of American geographers*, *44*(4), 315-348.

He, C. (2002). Information costs, agglomeration economies and the location of foreign direct investment in China1. Regional studies, 36(9), 1029-1036.

He, Z. L., Tong, T. W., Zhang, Y., & He, W. (2018). A database linking Chinese patents to China's census firms. *Scientific data*, *5*, 180042.

He, Z. L., Tong, T. W., Zhang, Y., & He, W. (2018). Constructing a Chinese patent database of listed firms in China: Descriptions, lessons, and insights. Journal of Economics & Management Strategy, 27(3), 579-606.

Head, K., & Mayer, T. (2004). Market potential and the location of Japanese investment in the European Union. *Review of Economics and Statistics*, *86*(4), 959-972.

Head, K., Ries, J., & Swenson, D. (1995). Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. *Journal of international economics*, *38*(3-4), 223-247.

Head, K., & Ries, J. (1996). Inter-city competition for foreign investment: static and dynamic effects of China's incentive areas. *Journal of Urban Economics*, *40*(1), 38-60.

Henderson, J. V., & Kuncoro, A. (1996). Industrial centralization in Indonesia. *The World Bank Economic Review*, *10*(3), 513-540.

Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial development in cities. *Journal of political economy*, *103*(5), 1067-1090.

Hou, Y. (2016). Traffic congestion, polycentricity, and intraurban firm location choices: a nested logit model for the Los Angeles metropolitan area. *Journal of Regional Science*, *56*(4), 683-716.

Jacobs, J. (1969). economy of cities.Random House, New York

Jaffe, A. B. (1989). Real effects of academic research. *American economic review*, *79*(5), 957-970.

Jean, R. J. B., Tan, D., & Sinkovics, R. R. (2011). Ethnic ties, location choice, and firm performance in foreign direct investment: A study of Taiwanese business groups FDI in China. *International Business Review*, *20*(6), 627-635.

Karaca‐Mandic, P., Norton, E. C., & Dowd, B. (2012). Interaction terms in nonlinear models. *Health services research*, *47*(1pt1), 255-274.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, *99*(3), 483-499.

Lai, H. C., Chiu, Y. C., & Leu, H. D. (2005). Innovation capacity comparison of China's information technology industrial clusters: The case of Shanghai, Kunshan, Shenzhen and Dongguan. *Technology Analysis & Strategic Management*, 17(3), 293-316.

Levinson, A. (1996). Environmental regulations and manufacturers' location choices: Evidence from the Census of Manufactures. *Journal of public Economics*, *62*(1-2), 5-29.

Liu, F., & Sun, Y. (2009). A comparison of the spatial distribution of innovative activities in China and the US. Technological Forecasting and Social Change, 76(6), 797-805.

Ma, R., & Liu, W. H. (2019). Spatial Concentration of Innovation Activities in China and the Role of Mega‐economic Zones. *China & World Economy*, *27*(2), 24-43.

MARSHALL, A. (1890). PRINCIPLES OF ECONOMICS, VOL. I.

- (1920). Principles of economics, an introductory volume.

Maskell, P., & Malmberg, A. (1999). Localised learning and industrial competitiveness. *Cambridge journal of economics*, *23*(2), 167-185.

McFadden, D. (1973) Conditional Logit Analysis of Qualitative Choice Behaviour,in Zarembka, P. (ed.) Frontiers in Econometrics, Academic Press, New York, 105-142.

McFadden, D. (2001). Economic choices. *American economic review*, *91*(3), 351-378.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of applied Econometrics*, *15*(5), 447-470.

Nie, H. H., Jiang, T., Zhang, Y. X.& Fang, M. Y. (2016). CHINA'S ZOMBIE FIRMS: CAUSE, CONSEQUENCE AND CURE. *China Social Sciences Press.*

OECD (1995). Science and Technology Indicators. Paris.

OECD. (2002). *The Measurement of Scientific and Technological Activities Frascati Manual 2002: Proposed Standard Practice for Surveys on Research and Experimental Development*. Cambridge University Press.

Pakes, A., & Griliches, Z. (1984). Patents and R&D at the firm level: a first look. In *R&D, patents, and productivity* (pp. 55-72). University of Chicago Press.

Polenske, K. R., & Ratanawaraha, A. (2007). Measuring the geography of innovation: a literature review, [w:] KR Polenske (red.). *The economic geography of in-novation, Cambridge University Press, Cambridge*, 30-59.

Poncet, S., Steingress, W., & Vandenbussche, H. (2010). Financial constraints in China: firm-level evidence. *China Economic Review*, *21*(3), 411-422.

Rasciute, S., Pentecost, E. J., & Marques, H. (2008). An empirical investigation of the determinants of the location of foreign direct investment in the Central and Eastern European countries using multilevel data.

Romijn, H., & Albaladejo, M. (2002). Determinants of innovation capability in small electronics and software firms in southeast England. *Research policy*, *31*(7), 1053-1067.

Shi, J. & He, C.F. (2018). Determinants of the spatial variation of new firm formation in China's metal product industry. *Geographical Research,* 37(7), 1281-1296.

Sigurdson, J., & Palonka, K. (2008). Innovative city in west China Chongqing. *European Institute of Japanese Studies Working Paper*, *239*.

Sirilli, G. (1998). Conceptualizing and measuring technological innovation.

Stein, J. C. (2003). Agency, information and corporate investment. In *Handbook of the Economics of Finance* (Vol. 1, pp. 111-165). Elsevier.

Sun, Y. (2000). Spatial distribution of patents in China. *Regional studies*, *34*(5), 441-454.

Sun, Y. (2003), Geographic patterns of industrial innovation in China during the 1990s. *Tijdschrift voor economische en sociale geografie*, 94: 376-389.

Tavassoli, S., & Carbonara, N. (2014). The role of knowledge variety and intensity for regional innovation. Small Business Economics, 43(2), 493-509.

Tianying, J. (2013). Study on spatial and temporal evolution and factors of regional innovation in China. *Economic Geography, 33*(6), 22-29.

Wei, S.H., Zhuo, J.J. & He, Y. (2011). The soatial distribution and changing tendency of regional innovation capability in China. *Science Research Management,* 32(4), 152-160.

Yang, C. H., Chiu, C. Y., & Tsou, M. W. (2017). Location Choice of Multinational and Local Firms in Vietnam: Birds of a Feather Flock Together? *The Japanese Economic Review*, *68*(1), 95-114.

Yi-xu, C. Y. Q. (2012). Measurable analysis on the difference changes of regional innovation capability based on the dynamic performance in china [J]. *China Population, Resources and Environment*, *3*.

Ze-Lei, X., Ru-Guo, F., & Xin-Ya, D. (2019). Measurement and Convergence of China's Regional Innovation Capability. *Science, Technology and Society*, *24*(1), 1–28.

Zhou, W.Y. & Xiang, Y. (2015) An empirical study on key factors of regional innovative capability about China's provinces and cities. Science Research Management. 36, 29-35.

Zheng, D., & Shi, M. (2018). Industrial land policy, firm heterogeneity and firm location choice: Evidence from China. *Land Use Policy*, *76*, 58-67

Zucker, L. G., Darby, M. R., & Brewer, M. B. (1994). *Intellectual capital and the birth of US biotechnology enterprises* (No. w4653). National Bureau of Economic Research.

Documents of policies on developing the National High-tech Industrial Development Zone. http://www.most.gov.cn/gxjscykfq/wj/. Retrieved Aprial 28, 2018