DIVERSITY OR SPECIALIZARION? UNDERSTANDING KNOWLEDGE SPILLOVER MECHANISMS IN CHINA

Shicong XU
PhD candidate, Department of Agricultural, Environmental and Development Economics, Suite 250 Ag. Administration Building, 2120 Fyffe Rd, The Ohio State University, Columbus OH 43210, USA
Shicong.x@gmail.com

Abdoul G. SAM
Professor, Department of Agricultural, Environmental and Development Economics, Suite 250 Ag. Administration Building, 2120 Fyffe Rd, The Ohio State University, Columbus OH 43210, USA,
Corresponding author
Sam.7@osu.edu

Abstract

China's rise to the top echelons of the world's economies was accompanied by an expeditious growth in domestic patent applications. Not surprisingly, this phenomenon has spawned a growing literature trying to sort out the determinants of patented research in China. However, mostly due to data limitations, many of the papers on this topic use aggregated innovation data at the industry, prefecture, or province levels. In this paper, we examine the empirical validity of important theories of knowledge spillover in the context of China at a micro-level, using a firm-level panel dataset comprised of publicly traded companies listed in the Shanghai and ShenZhen Stock Exchanges during the 2006-2010 period. Our study sheds light on whether locating near innovative firms increases patenting activity in general, regardless of the industry membership of these neighboring firms. We also explore how industry makeup, measured by the number of firms in the same or different industries, affects firm-level patenting activity. Our econometric results show that the number of patent applications by firms in close geographic proximity of a firm of interest has a significant and positive impact on that firm's successful patent applications. In addition, we find that proximity to firms in the same industry reduces innovation while locating near firms from different industries stimulates innovation.

Keywords: patents, knowledge diffusion, MAR spillover, Jacobs spillover, China
JEL classification: O31, O32, O33, R12, D22

1. Introduction

China's swift rise as a global economic power in recent decades has been associated with an equally impressive increase in the number of patented innovations. The number of granted patents reportedly rose from 138 in 1985 to over 200,000 in 2005, to more than 800,000 in 2010 (China Statistical Yearbook of Science and Technology 2019; Yueh 2009). Consequently, several studies have sought to understand the mechanisms of technological innovation in China. These papers have mostly focused on the effects of trade and foreign direct investment (FDI), research and development (R&D), skilled labor, and type of firm ownership. We empirically explore the impact of inter-firm knowledge flows among neighboring firms, using data from a broad set of industries in China. Our focus on firm-to-firm knowledge spillovers in China is premised on the Marshallian and Jacob spillover theories that the spatial clustering of firms from the same or different industries can prompt innovation (Xie et al. 2019; Zhu et al. 2019). Most of the studies investigating knowledge spillover in China, use higher levels of data aggregation, such as province, prefecture/city-level data, which constitutes an important limitation because the flow of knowledge permeates across border lines via multiple modes of diffusion including from firm to firm. Also, the level of interdependence among firms within a given special scale may dictate the degree and extent of knowledge spillover (Glaeser et al. 1992).

We study knowledge diffusion in China at a more micro-level, using firm-level data, with the goal of answering two specific research questions. The first research question explores the effect of pure spatial proximity between firms on knowledge spillover, regardless of industry composition. Knowledge spillover can create a snowball effect that increases overall productivity, product quality, operational efficiency, and market competitiveness within a region synergistically. The second research question investigates the effect of industry makeup on patent applications at the firm level within a given area. The rest of this study is organized as follows. The second section provides a background and summarizes the existing literature. The third section discusses the data and data sources used in this study. The fourth section outlines the empirical methods. The fifth section presents our results, and the sixth section concludes.

2. Literature Review

2.1. MAR and Jacobs knowledge spillovers

The creation and spread of knowledge have been popular topics for a long time. A number of studies in developed economies (mainly US, Japan, and the European countries) have found evidence that geographic proximity drives knowledge diffusion (e.g. Jaffe 1989; Audretsch and Feldman 1996a; Aldieri 2011; Bottazzi and Peri 2003) in support of two leading theories of intellectual spillovers. The Marshall-Arrow-Romer (MAR) theory of spillover argues that firms in the same industry will experience a higher knowledge spillover if they are located close to each other. The Jacobs spillover theory, on the other hand, argues that firms in different industries will experience a higher level of knowledge spillover if they are located close to each other. The Jacobs theory specifically argues that firms within *complementary* industries benefit from knowledge spillover. It is harder to measure complementary relationships between firms. In our study, we simply consider the number of firms that are *not* in the same industry as a measurement for Jacobs spillover. As more firms enter the complementary supply chain, more knowledge spillover will occur. In other words, MAR focuses on common industry knowledge spillover while Jacobs focuses on diversity-driven knowledge spillover. MAR spillover is often referred to as specialization spillover and the Jacobs spillover as diversity spillover. The two types of spillovers are not mutually exclusive; it is possible that both types of knowledge diffusion can be observed at the same time. The evidence in support of one or the other theory is dependent, in part, on how industry and geographic clusters are specified and how knowledge spillover is defined and measured: economic growth indicators, productivity measures, or innovation indicators (Beaudry and Schiffauerova 2009). Jacobs spillover was frequently detected when industry specifications are at a medium level of detail (represented by 3-digit industry code). On the other hand, there seems to be higher level of MAR detection when industries are more broadly defined (at 1- or 2-digit level). This sort of difference in effects is also observed across industry types: high-tech, medium-tech, low-tech, and service industries. Across these different types of industries, MAR spillover was observed to be significant often in low- and medium-tech industries while Jacobs spillover was observed to be significant in high-tech industries. Service industries showed the lowest evidence of significance in both the MAR and Jacobs spillovers. The next most common source of heterogeneity in spillover type is associated with how the geographic unit is defined – from the broadest definition of a geographic unit being at the state/provincial level to the smaller units being at the counties, labor zones, zip-code levels, and MSA levels. The probability of detection of both MAR and Jacobs spillover increases as the geographic unit specifications become smaller.

2.2. Determinants of innovation in China

Several studies have tried to explain the patent boom mentioned above. In particular, FDI and trade have been found to significantly correlate with knowledge spillover in China at the province, city, and industry levels (e.g., Ning et al 2016; Fu 2008; Liu and Buck 2007). FDI serves as a conduit of knowledge transfer from multinational corporations to their local subsidiaries and it crowds in financial resources, new technologies, and skilled laborers (Fu 2008; Chuang and Hsu 2006; Hu and Jefferson 2009; Liu 2002; Yueh 2009; Yang and Lin 2011; Gao 2004; Huang and Wu 2012; Shang et al. 2012; Ning et al. 2016; Lin et al. 2009). Per Eun et al. (2006), FDI-driven knowledge spillover is commonly observed and a normal phenomenon in fast developing countries like China. While FDI gives domestic firms access to new knowledge, many studies have found that industries and regions must muster a minimum level of innovation absorbability before any knowledge spillover can be observed (Huang et al. 2010; Fu 2008; Yang and Lin, 2011; Liu and Buck, 2007). Some studies have found no relationship between FDI and innovation (Shang et al., 2012) or a negative one (Liu et al. 2010). Liu and et al (2010) argue that increased FDI flows to local firms in the same industry may heighten competition which, in turn, dampens firms' appetite for risky forms of investments such as innovation. Their analysis is, however, confined to a sample of firms in Beijing's Zhongguancun Science Park over the period 2000-2003.

Researchers have found a positive and significant spillover effect from university and research institutions to industries (e.g., Shang et al., 2012). Per Acs et al (1994), the innovation spur from public research institutions is mostly captured by smaller firms; larger firms were found to benefit mostly from their own R&D investment (Acs et al. 1994). These findings are intuitive in the sense that smaller firms are generally unable to undertake their own innovation and therefore tend to rely more heavily on knowledge generated from the public sector and research institutions. Larger firms on the other hand may be deep-pocketed enough to carry out their own R&D investment on innovation. For China, however, because of the large gap between public and private innovation research, some studies have argued that the spillover between the two may not be as significant as has been found in developed countries. Huang and Wu (2012) indicate that most of the technology patents are put forth by universities and public research institutions and that, in the case of nanotechnology, firms in China have not been able to capitalize on R&D from public research institutions.

Several resource-related characteristics such as R&D, skilled labor, and capital intensity have also been found to be critical inputs to the innovation production function (Lucking et al., 2018; Ning et al. 2016; Huang and Wu 2011; Yueh 2009; Yang and Lin 2011; Rho and Moon 2014; Chang and Sam 2015; Carrion-Flores et al. 2013). More skilled workers are found to have a positive and significant effect on innovation. For example, Shang et al. (2012) find that skilled workers have a more pronounced effect on innovation for research institutions than for firms. They speculate that the reason behind this observation is that research institutions are more interested in novel technological directions (inventor patents) which require more highly skilled workers. Firms, on the other hand, are more interested in perfecting existing processes and products, hence tend to focus on less innovative (design and utility) patent types which require fewer skilled workers.

More closely related to our work are the studies of Li et al. (2019) and Wang and Wu (2016). Li et al. (2019) examine the spillover effect of university research on firm-level innovation, measured by patents. Like our approach, the distances between firm headquarters and nearest research universities are calculated using ArcGIS. Their results show that patent applications by firms are inversely related to their distance to the nearest university and positively related to the research capability of nearby universities. Our research focuses on firm-to-firm rather than university-to-firm knowledge diffusion. Wang and Wu (2016) find evidence that "geographical proximity plays a significant role in knowledge spillover and diffusion" among domestic Chinese firms. However, this study is confined to only one industry (Electronics) and covers one year (2009). Also, in the same vein of research as ours, Rho and Moon (2014), Wang et al (2014), and Shang et al. (2012) empirically study spatial dependence in province-level innovation in China using different types of spatial weight matrices to account for distance between provinces. Rho and Moon (2014) find evidence of interdependence in regional innovation up to a distance of 2,000 kilometers between provincial capitals. Unlike the studies above which, for the most part, are based on regional or industry-level data or are limited to a particular sector, we analyze knowledge flows between neighboring firms from a broad panel of industries in China. As such, our study significantly adds to the important body of literature on knowledge spillover in China.

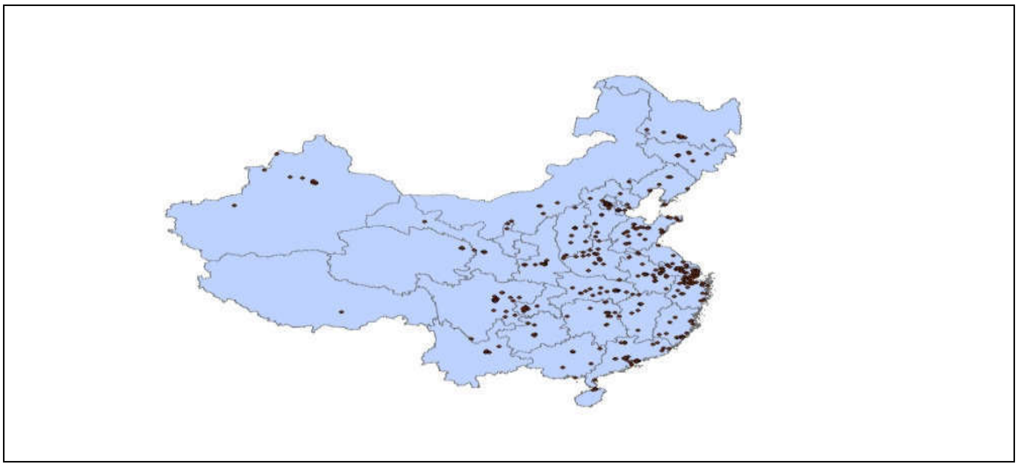
3. Data

We collected and aggregated our data from several sources. This resulted in an unbalanced panel of 1,654 firms that are listed either on the Shanghai Stock Exchange or the ShenZhen Stock Exchange for the years 2006-2010 during which data on all variables was available. Fifteen industries are represented in our data (Table A1 in appendix). Our dependent variable (patents) was obtained from the China National Intellectual Property Administration

(CNIPA), along with the addresses of the patent assignee names. Chinese patents can be divided into three main types: invention, utility, and design. Invention patents are similar to utility patents in the US. Invention patents must demonstrate precedence of discovery or meaningful improvement of a previous process or product innovation. The protection length for invention patents is 20 years whereas for utility and design patents it is 10 years (Moga, 2017). Because utility and design types of patents do not go through an examination process and are not considered to be *innovative* ideas per se, we only consider the invention patents which constitute about half of all patents.

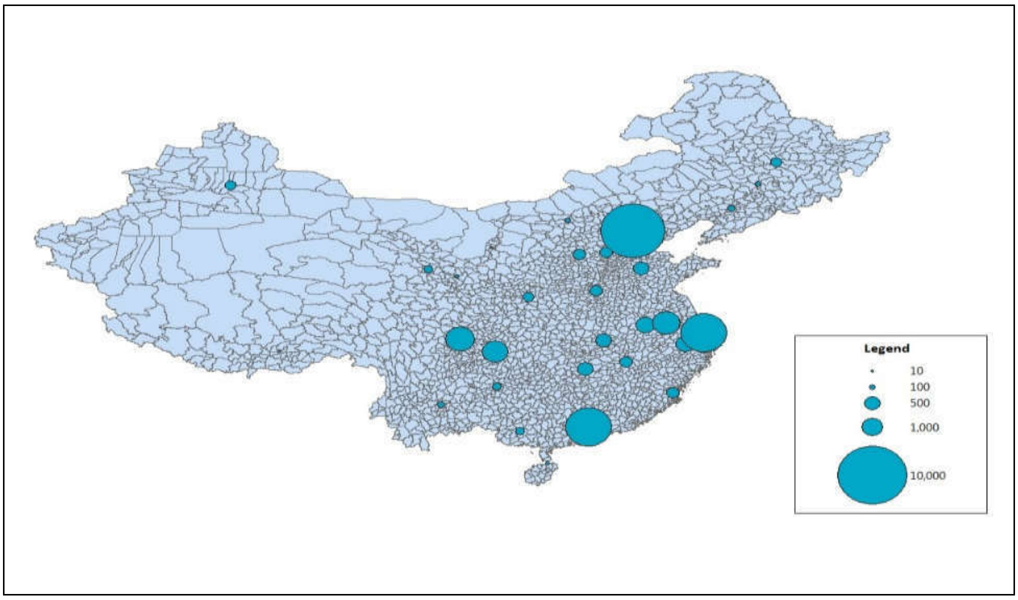
Figures 1 and 2 display, respectively, the spatial distribution of firms in our sample and matched patents in 2010.

Figure 1: Spatial distribution of sample firms



Source: Shanghai Stock Exchange or the ShenZhen Stock Exchange

Figure 2: Spatial distribution of patents



Source: China National Intellectual Property Administration

The figures show that there are a few spatial clusters where we observe most firms and patents: northeastern region, southern coastal region, and central/east coastal region. These clusters are the locations of three megacities in China: Beijing (northeast), Shanghai (central/east coastal), and ShenZhen (south/coastal). The largest science parks are also found in these three regions; we use fixed effects to accounts for the science parks. Patents were matched to each firm based on information reported in the patent application. One potential limitation of our analysis is that it assumes that all innovation within a firm is taking place at the recorded patent assignee/firm address which may not be the case. It is possible that some of the firms have subsidiary facilities, located distantly from the assignee address, that generate patentable innovation. Unfortunately, our data is not detailed enough to identify the exact location where the innovation is actually taking place. More information regarding matching methods and accuracy check can be found in the paper published by He et al., (2018). He et al., (2018) published a paper on the method used and accuracy check regarding the matching of patents to companies that own them. The authors also made the matched patent data publicly available to other researchers. We obtained all firm level data from the WIND Financial Terminal (i.e. R&D expenditure, capital expenditure, sales, profit, type of ownership, number of workers, average inventor per patent). We calculated R&D intensity as firm level R&D spending divided by total sales, and capital intensity as capital expenditure divided by total sales. We take the log of profit and number of workers due to the skewness and large range of these two variables across different firms. We control for state ownership and foreign ownership. State ownership was computed as the share of stock owned by the State divided by total number of outstanding stocks; foreign ownership was calculated likewise as the share of stocks owned by foreigners. We obtained all prefecture level variables from the China Economic and Social Development Statistics Database. Provincial data were retrieved from National Bureau of Statistics of China (NBS).

Tables 1 and 2 presents the variable definitions and summary statistics of the data, respectively. Several data constraints lead to the limitation of the study sample to the period 2006-10 and prevented us from directly controlling for FDI and skilled workers in our estimations. First, data for firm-level R&D expenditure is not available until 2006. The need to control for R&D, one of the most important inputs to the generation of knowledge, explains the beginning of the sample period in 2006. Second, our dependent variable, matched patent data, ends in year 2010. With respect to FDI data, it was mostly retrieved from the China Statistical Yearbook and provincial statistical yearbooks in the early 1990s. However, since 1996, only FDI data for larger firms and from select industries is available to researchers (Zhou et al. 2002; Hericourt and Poncet 2009; Xu and Sheng, 2011). Firm-level FDI data collected and made available by the World Bank Investment Climate Surveywas limited to the period of 1999-2003 and was only done in selected cities and provinces (Hericourt and Poncet 2009; World Bank 2003). In other words, the available FDI firm level data falls out of the range of our study period and does not encompass all regions. While we do not have a firm-level FDI control, we believe that the inclusion provincial and science park fixed effects in our regressions should mitigate the problem to a large extent since the studies at a more aggregated level indicate that FDI varies across regions with flows mostly coastal provinces and megacities (Su and Jefferson, 2012; Tanimoune et al., 2013; Zhang and Roelfsema, 2014).

Detailed information at the firm level is very limited in the case of China which explains why current studies that have skilled workers as a control are either at an aggregated level (Ning and Li 2016; Shang et al 2012; Fu 2008) or only use select industries (Liu et al 2010). Liu et al (2010) use the presence of returnee entrepreneurs as a proxy for skilled labor among a set of high-tech industries in Beijing's Zhongguancun Science Park over the period 2000-2003. For lack of data on skilled workforce at the firm level, we include the number of workers and the average number of inventors per patent by firm in our regressions. The unique Hukou System in China can also serve as a proxy for skilled labor. Hukou is the internal passport system that restrict free labor flows within the country. Unlike most of the population whose residence is restricted to their birthplace, highly educated Chinese citizens tend to experience little to no internal migration restrictions (Wu and Treiman 2007; Chan and Zhang 1999; Fu and Gabriel 2012). In this case, we are referring to movements from a lower tier location to a higher tier location such as rural to urban, small city to medium/large size cities, or medium to large cities. Generally, the barrier of migration can be lowered when the expected contribution of an individual is high. The effect of the Hukou system on internal migration in China has been studied heavily by many researchers.

Table 1: Variable Definitions

Variable	Variable Description
inventor	Number of successful inventor patent applications submitted by firm i at time t
ninventor5	Total inventor applications submitted by neighboring firms within 5km of radius at time t around firm i

ninventor20	Total inventor applications submitted by neighboring firms within 20km of radius at time *t* around firm *i*
ninventor35	Total inventor applications submitted by neighboring firms within 35km of radius at time *t* around firm *i*
ninventor50	Total inventor applications submitted by neighboring firms within 50km of radius at time *t* around firm *i*
ninventor100	Total inventor applications submitted by neighboring firms within 100km of radius at time *t* around firm *i*
ninventor200	Total inventor applications submitted by neighboring firms within 200km of radius at time *t* around firm *i*
ninventor300	Total inventor applications submitted by neighboring firms within 300km of radius at time *t* around firm *i*
marlevel5	Number of firms in the same industry within 5km of radius at time *t* around firm *i*
marlevel20	Number of firms in the same industry within 20km of radius at time *t* around firm *i*
marlevel35	Number of firms in the same industry within 35km of radius at time *t* around firm *i*
marlevel50	Number of firms in the same industry within 50km of radius at time *t* around firm *i*
marlevel100	Number of firms in the same industry within 100km of radius at time *t* around firm *i*
marlevel200	Number of firms in the same industry within 200km of radius at time *t* around firm *i*
marlevel300	Number of firms in the same industry within 300km of radius at time *t* around firm *i*
R&D Intensity	R&D Intensity = R&D expenditure/Sales. To get percentage points R&D intensity is multiplied by 100.
State Ownership	State Ownership = Number of State-owned stocks/total issued stocks
Foreign Ownership	Foreign Ownership = Number of Foreign-owned stocks/total
Capital Intensity	Capital Intensity = Capital expenditure/Sales
Profit	Total amount of profit measured in 10000 observed by firm *i* at time *t*
Workers	Total number of workers firm *i* has at time *t*
Capital Expenditure	Total amount of capital expenditure by firm *i* at time *t*
Average Inventor	Average number of inventors per patent by firm *i* at time *t*

We will not go into details on this specific topic since it is not the focus of our research. The migration patterns for more educated groups tend to be toward larger cities or megacities where there is greater concentration of human capital, better local amenities, and employment opportunities (Fu and Gabriel 2012; Fan 2002). Hence, the combination of the average number of inventors per patent, science park and regional fixed effects will capture the effects of skilled labor on firm innovation in our models. It should also be noted that, while skilled workers are a main input of innovation for research institutions, their effect on indigenous firms is unclear in the case of China (Shang et al., 2012) as discussed in the previous section.

4. Methods

We follow a two-step method similar to that outlined by Wallsten (2001). First, we used ArcGIS software to plot each firm's location using their address from the CNIPA database, latitude, and longitude coordinates. Once we pinpoint the location of a firm, we use a series of radii to draw a circular-shaped-buffer around it in such a way that it is the center of the circle. The series of radii we used in our study are 5, 20, 25, 50, 100, 200, and 300 kilometers (km). Key variables are derived within each buffer (i.e., control for total number of firms, total number of patents, total R&D spending, number of workers, and number of firms in the same industry as centering firm *i* with respect to each increase in the radius). This sequential widening of the buffer zone allows us to investigate how the effects of the variables change marginally as distance to the centering firm is increased.

Table 2: Summary Statistics

Variable	Mean	Std. Dev.
Inventor	35.018	281.468
ninventor5	129.141	666.538
ninventor20	458.209	1155.714
ninventor35	551.212	1265.729

ninventor50	575.367	1280.325
ninventor100	723.840	1431.437
ninventor200	1062.843	1538.448
ninventor300	1333.045	1599.353
marlevel5	1.692	1.265
marlevel20	4.020	4.923
marlevel35	5.297	6.839
marlevel50	5.973	7.332
marlevel100	8.790	9.496
marlevel200	18.095	19.206
marlevel300	27.247	27.079
R&D Intensity (percentage points)	1.355	4.154
State Ownership	0.197	0.229
Foreign Ownership	0.01986	0.0845
Capital Intensity	0.092	0.147
Profit (CNY 1,000)	609,000	4,860,000
Capital Expenditure (CNY 1,000)	1,070,000	7,850
Workers	8,177.415	28,760.91
Average Inventor	3.424	2.399

Once we obtained our key variables within each buffer, we applied a panel data method to estimate how spatial proximity and industry makeup within an assigned geographic boundary affect knowledge spillover in China at the firm level. Specifically, we explore our research questions using the reduced form model:

$$P_{itr} = \alpha_r + \beta_r Pat_{itr} + \delta_r C_{itr} + \theta_r X_{i\,tr} + \tau_r Per_{ir} + \lambda_{tr} + \varphi_{ir} + \rho_{ir} + s_{ir} + \varepsilon_{i\,tr} \tag{1}$$

where the radius r = 5, 20, 35, 50, 100, 200, 300 km, the dependent variable, P_{itr} , is the number of successful patent applications submitted by firm i at time t located in the buffer with radius r. Pat_{itr} is the patent count within radius r around firm i at time t. Our key variable of interest and other control variables within each buffer will change as the area changes. C_{itr} represents the number of firms within the same industry as the centering firm i and captures the effect of industry makeup within radius r. X_{itr} are the firm level controls; Per_{ir} are the prefecture-level controls; λ_{tr} are the time fixed effects; φ_{ir} are the industry fixed effects; ρ_{ir} are the provincial fixed effects; s_{ir} are the science park fixed effects; and α_r , β_r, δ_r,θ_r, τ_r are parameters to be estimated with β_r and δ_r being the main parameters of interest in our reduced form model.

The key variables we control for at the firm level are the number of workers, R&D intensity, capital intensity, profit, ownership information, and the average number of inventors per patent. These control variables are chosen based on previous work as discussed in the literature review. We control for firm profits because economists have argued that a firm's internal resources are an important determinant innovation due to asymmetric information related to innovation research. That is, innovators have more information about the likelihood of success of a research project and the risk of intellectual property infringement precludes full disclosure to potential investors. Such asymmetric information may lead investors to demand a premium in order to finance research projects (Hall 2002) making external financing more expensive than internal financing (profits). We include two interaction terms (average number of inventors*R&D and average number of inventors*capital intensity) in order to capture the synergistic effects between innovation inputs within firms. We include prefecture GDP in order to control for the possibility that firms are sorted into an area due to the first two drivers of Marshallian agglomeration that were discussed earlier: decrease in the cost of goods (up and down the supply chain) and workers. The reason, prefecture GDP can capture the value generated by the local inputs such as establishments and size of labor pool in each prefecture. We control for regional fixed effects in our estimations to capture time-stable heterogeneity across regions such as provincial policies that may affect firm patenting behavior. We consider a total of 31 regions: 22 provinces: Anhui, Fujian, Gansu, Guangdong, Guizhou, Hainan, Hebei, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Jiangxi, Jilin, Liaoning, Qinghai, Shaanxi, Shandong, Shanxi, Sichuan, Yunnan, and Zhejiang; 4 major municipalities: Beijing, Chongqing, Shanghai, and Tianjin; and 5 autonomous regions: Guangxi, Inner Mongolia, Ningxia, Tibet, and Xinjiang. Science park dummy variables control for the effect of high concentrations of innovative firms within a small area thanks to policies of the central government (Walcott, 2002). Lastly, industry and time fixed effects are added as controls for any time trends and heterogeneity across industries.

We ran a series of regressions using key variables calculated for each buffer. The first regression was ran using a 5km radius. The second regression was ran using a 20 km radius, while controlling for the data from the 5km of radius. In other words, the second regression contains patent data from both the 5km and 20km radius buffers. Controlling for the smaller buffer allows us to see the incremental effect of neighboring patents beyond the first buffer on centering firm *i*'s patent activity. The next few regressions were ran using 35, 50, 100, 200, and 300 km of radius utilizing the same approach. For each additional regression, we expand to the next buffer size while controlling for the smaller areas to estimate the incremental effect of patent counts from the additional neighboring firms.

One significant challenge for our regression model is the presence of endogeneity with respect to neighboring patent count. Our dependent variable is the number of successful patent applications submitted at time *t* by firm *i.* Our first covariate of interest is the total number of successful patent applications submitted by all other firms within radius *r* at time *t* around centering firm *i*. Theoretically, if we expect that centering firm *i*'s patenting behavior is influenced by neighboring firms' patent applications then, according to this logic, we should expect that the other direction of the relationship is also true. That is firm *i*'s patenting behavior should, in turn, also influence neighboring firms' patenting activities. We rely on an instrumental variable approach (IV) to address this reverse causality. We use two variables to serve as our instruments: workforce of neighboring firms and R&D expenditure by neighboring firms around firm *i* at time *t* within radius *r*. Both R&D and workers are important inputs that produce patentable innovation. Thus, we expect the instruments to be highly correlated with the endogenous variable. At the same time, other firms' inputs (R&D spending and size of workforce) should not directly affect the centering firm's innovative activity. One may argue that the centering firm could alter its innovation policy after observing neighboring firms' input decisions. While it is possible that firm *i* may increase its own innovation input (such as investing more in R&D or delegating more workers toward innovation) after observing neighboring firms' R&D decisions and workforce changes, this represents an indirect effect through the endogenous variable. Therefore, it is reasonable to conclude that the instruments used do not directly explain the dependent variable. Test of relevance and exogeneity are presented and discussed later in the paper.

5. Results and Discussion

Our main regression results can be found in Tables 3 and 4. In Table 3, we present our results for radius stacking in order to answer our first research question. Table 4 presents the results for progressively larger radii but without controlling for the smaller areas each time a larger radius is used. Regression results from Table 3 answer our second research question by including the number of firms that are in the same industry as centering firm *i* within radius *r* at time *t*; this serves as our MAR spillover indicator. Table 3 column 1 presents our first regression results for the 5km radius around each firm.

Table 3: Regressions for Progressively Larger Firm Neighborhoods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	**R=5**	**R=20**	**R=35**	**R=50**	**R=100**	**R=200**	**R=300**
ninventor300							-0.000571
							(0.0261)
ninventor200						0.00865	0.00878
						(0.0217)	(0.0294)
ninventor100					-0.0328	-0.0393	-0.0395
					(0.0697)	(0.0729)	(0.0720)
ninventor50				0.0871	0.146	0.140	0.141
				(0.179)	(0.235)	(0.234)	(0.233)
ninventor35			-0.0159	-0.113	-0.140	-0.135	-0.136
			(0.0332)	(0.204)	(0.222)	(0.221)	(0.221)
ninventor20		-0.0142	-0.00112	0.0106	0.0119	0.0106	0.0107
		(0.0258)	(0.0383)	(0.0461)	(0.0469)	(0.0467)	(0.0467)
ninventor5	0.118*	0.130*	0.116*	0.116*	0.121*	0.120*	0.118*
	(0.0672)	(0.0673)	(0.0679)	(0.0678)	(0.0678)	(0.0675)	(0.0672)
R&D Intensity$_{t-1}$	10.04**	8.534**	10.31***	10.19***	10.18***	10.01**	10.04**
	(3.944)	(3.810)	(3.914)	(3.910)	(3.931)	(3.936)	(3.944)
Log (Workers)	45.59***	41.53***	44.49***	44.57***	45.30***	45.52***	45.59***
	(10.17)	(10.21)	(9.808)	(9.845)	(10.07)	(10.11)	(10.17)
Log (Profit)	0.410	0.469	0.448	0.450	0.414	0.414	0.410
	(0.962)	(0.922)	(0.956)	(0.954)	(0.963)	(0.960)	(0.962)
Capital Intensity	-30.27	-28.27	-26.66	-30.23	-30.68	-30.26	-30.27
	(96.78)	(94.05)	(95.35)	(95.62)	(96.17)	(95.99)	(96.78)
State Ownership	-38.73	-28.61	-33.32	-32.47	-37.59	-38.48	-38.73
	(35.58)	(32.49)	(33.28)	(33.27)	(35.29)	(35.38)	(35.58)
Foreign Ownership	-138.6	-118.4	-138.5	-154.9	-134.7	-137.2	-138.6
	(184.7)	(187.5)	(176.9)	(181.6)	(185.2)	(185.4)	(184.7)
Average Inventor	-2.135	-1.572	-1.894	-2.045	-2.103	-2.129	-2.135
	(3.483)	(3.318)	(3.425)	(3.439)	(3.461)	(3.455)	(3.483)
Average Inventor * R&D Intensity$_{t-1}$	0.569	0.547	0.554	0.594	0.583	0.576	0.569
	(0.917)	(0.882)	(0.912)	(0.914)	(0.919)	(0.917)	(0.917)
Average Inventor * Capital Intensity	8.118	8.139	7.930	8.694	8.073	8.133	8.118
	(21.59)	(20.81)	(21.33)	(21.37)	(21.52)	(21.47)	(21.59)
Prefecture GDP	0.00123	-0.000420	0.000664	0.000565	0.00126	0.00119	0.00123
	(0.00447)	(0.00416)	(0.00418)	(0.00419)	(0.00448)	(0.00447)	(0.00447)
Observations	879	879	879	879	879	879	879

Notes: all regressions include year, firm, industry, province/region, and science park fixed effects. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The key variable of interest, ninventor5, indicates how total neighboring patents within a 5km of radius affect the centering firm patenting behavior. Our results show that an increase of 10 patent applications collectively by the neighboring firms (about 7.75% of the sample mean--see Table 2) increases center firm's patent applications by about 1.2 patents (about 3.4% of the sample mean) on average. In regression 2, we increase the radius to 20km. Again, we control for the effect on innovation by neighboring firms within 5km of radius. Controlling for the smaller radius allows us to investigate the incremental effect of the increasing distance--above and beyond the effect of innovation of neighboring firms within the 5 km radius. Hence, the marginal effect reveals the effect of neighboring patents on firm i's patenting behavior from 5km – 20km. We find that the neighboring firms' patenting activities have a positive and significant effect within radius of 5km, but the effect is no longer significant outside of the 5km radius. This observation remains when we increase the radius to 35, 50, 100, 200, and 300km and indicates that, while firms benefit from spatial knowledge spillover, the effect does not improve with increasing in distance. Knowledge spillover could not be detected outside of 5km of radius in that all the incremental marginal effects outside of the 5km are insignificant.

Our results can be counterintuitive at first glance. As we increase the radius, a bigger area is being considered; therefore, more firms and more patents would be observed within the area. The knowledge spillover literature has suggested that more neighboring patents should increase other firms' patent applications (e.g. Jaffe et al.1986; Aldieri 2011; Bottazzi and Peri 2003). However, many of these studies did not consider how an incremental increase in distance between firms affects innovation activities. Our results have shown that while there is knowledge spillover, it's effect in terms of stimulating patentable research is very much limited by geographic proximity in China. Most of the studies that were done in developed countries have found the effect of knowledge spillover diffuses to a larger scale. For example, Bottazzi and Peri (2003) studied knowledge spillover in European regions and found that knowledge spillover diffuses up to a 300km radius. Audretsch and Feldman's (1996b) findings about the interaction between knowledge spillover and industry life cycle can provide a possible explanation for the limited geographical reach of knowledge spillover found in our results. They argue that knowledge spillover spurs innovation within a small area at the beginning of an industry's cycle because the lack of accepted standards and a high degree of product uncertainty during the early stages of an industry's life cycle make it harder for firms to know what consumers desire without proximity to knowledge sources. In other words, "tacit knowledge" plays an important role in the early stage of industrial development and it requires physical proximity to obtain it. Once an industry matures, the uncertainty about product standards and consumer preferences are much reduced, hence finished products become a springboard to spread knowledge farther geographically. China is still going through the process of economic transformation and has yet to reach full economic maturity, which could explain our empirical result that the area of effect in terms of knowledge spillover is significantly smaller than in developed economies. Our results suggest that the current studies of knowledge spillover in China that rely on province level data are too broadly aggregated.

Table 4: MAR Regressions

Variables	(1) R=5	(2) R=20	(3) R=35	(4) R=50	(5) R=100	(6) R=200	(7) R=300
ninventor	0.182**	0.0350	-0.0172	-0.0144	-0.00531	-0.0102	-0.0162
	(0.0793)	(0.0345)	(0.0294)	(0.0273)	(0.0237)	(0.0202)	(0.0194)
marlevel	-12.48	-9.582***	-5.343**	-5.286**	-3.184*	-1.335	-0.707
	(11.25)	(2.943)	(2.336)	(2.173)	(1.673)	(0.937)	(0.713)
R&D Intensity$_{t-1}$	8.584**	7.812**	10.27***	9.615***	8.412**	7.982**	8.672**
	(4.018)	(3.580)	(3.700)	(3.658)	(3.573)	(3.533)	(3.576)
Log (Workers)	40.75***	40.54***	43.65***	42.79***	41.54***	40.58***	42.73***
	(10.75)	(9.722)	(9.143)	(9.230)	(9.489)	(9.640)	(9.494)
Log (Profit)	0.544	0.474	0.332	0.356	0.374	0.327	0.278
	(0.971)	(0.865)	(0.906)	(0.894)	(0.871)	(0.854)	(0.873)
Capital Intensity	-34.01	-41.56	-28.15	-26.44	-17.67	-35.00	-40.76
	(98.22)	(88.73)	(89.92)	(89.34)	(88.34)	(87.24)	(87.59)
State Ownership	-26.08	-34.11	-40.08	-38.33	-35.92	-34.84	-37.83
	(34.26)	(30.35)	(31.21)	(30.90)	(30.49)	(29.88)	(30.26)
Foreign Ownership	-111.5	-262.5	-245.2	-233.9	-229.7	-210.5	-212.8
	(196.2)	(173.6)	(155.0)	(157.7)	(167.4)	(172.6)	(166.9)
Average Inventor	-1.879	-1.090	-1.227	-1.012	-1.113	-1.035	-1.286
	(3.486)	(3.123)	(3.229)	(3.191)	(3.106)	(3.059)	(3.094)
Average Inventor* R&D Intensity$_{t-1}$	0.928	0.624	0.279	0.284	0.288	0.308	0.296
	(0.929)	(0.837)	(0.856)	(0.844)	(0.820)	(0.803)	(0.811)
Average Inventor* Capital Intensity	8.685	7.081	5.660	5.131	4.800	7.915	8.886
	(21.82)	(19.54)	(20.09)	(19.90)	(19.51)	(19.22)	(19.38)
Prefecture GDP	0.00145	0.00362	0.00331	0.00274	0.00175	-0.000344	-0.000376
	(0.00441)	(0.00420)	(0.00418)	(0.00417)	(0.00420)	(0.00400)	(0.00388)
Observations	879	879	879	879	879	879	879

Notes: all regressions include year, firm, industry, province/region, and science park fixed effects. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4 allows us to respond to our second research question pertaining to the interaction between knowledge spillover, industry makeup, and geographic proximity. In this table, we no longer focus on the incremental marginal effect with the increase in radius; instead, we show the total effect within a radius. In specification (1) of Table 4, our key explanatory variable (ninventor--total inventor applications submitted by neighboring firms within radius r) is still significant and positive, and the magnitude of the effect is larger than in Table 3: an increase of 10 neighboring patent applications leads to an increase of about 1.8 patent applications for the centering firm, corresponding to about 5% of the sample mean, on average. This result is only significant and positive within a 5km radius.

The MAR effect (coefficient on marlevel) is not statistically significant for the 5 km radius indicating that co-locating in close proximity with firms of the same line of business does not adversely affect the innovation spur. Given the definition of MAR spillover (number of firms that are in the same industry as the centering firm), the way we construct the Jacobs spillover variable (number of firms that are not in the same industry), and since we are controlling for the total number of patent applications by neighboring firms, the coefficient for Jacobs spillover is simply the negative of the MAR (marlevel) coefficient in Table 4. For this reason, the equivalent of Table 4 for the Jacobs spillover measure (instead of MAR spillover) is omitted. Given the size of the coefficient (-12.48), the non-significance of the MAR effect could also be due to lack of statistical power because of the small number of firms. As the radius is increased, we observe a significant and negative effect of MAR spillover (conversely a positive and significant effect of Jacobs spillover) on firm i’s patenting behavior in most of the specifications. In specification (2) for example which corresponds to a radius of 20 kms, we find that one additional firm within the same industry decreases the centering firm’s patent applications by nearly 9.6 units. This negative effect remains significant as radius increases; however, the magnitude of effect decreases rapidly. Beyond a100km radius, industry makeup is no longer a significant factor that affect firms’ patenting behavior.

Together, Tables 3 and 4 indicate the existence of knowledge spillover among Chinese firms; however, this effect is very limited to geographic proximity--5kms--and depends on industry makeup. Concentration of firms in the same line of business within a 100km of radius is detrimental to firm level innovation while having a variety of firms from different industries promotes innovation. Our findings lend support to previous studies that firm diversity is key to innovation stimulus (e.g., Beaudry 2009). We conjecture that the uncertainty from increased competition between alike firms leads to a decrease in high-risk investments such as patentable research. Lucking et al. (2018) find that R&D by product market competitors reduces patenting activity of a firm because of business stealing by the successful innovator. They argue that R&D investments by competitors reduce the marginal benefit of R&D thereby lowering a firm’s incentive to invest in R&D itself.

To gauge the validity of our IVs, we present the first stage regression in appendix (Table A2 in appendix). Both instruments are highly significant, indicating relevance. We also present results of the Hansen overidentification test to check on the validity of the exogeneity assumption. The null hypothesis of the Hansen test is that the instruments are uncorrelated with the error term in the main regressions (equation 1). All of our p-values are larger than 10%. In summary, both tests of relevance and exogeneity substantiate our argument of instrument validity.

6. Conclusion

In this study, we investigate knowledge spillover in China using firm-level panel data and ArcGIS to create uniform spatial scales around each firm. Our results provide important insights pertaining to well-known theories of knowledge spillover. First, geographic proximity between firms increases knowledge spillover, a finding that is consistent with previous studies in this realm (i.e. Shang et al. 2012; Aldieri 2011; Glaeser et al. 1992). Second, locating close to firms in the same line of business (indicated by the number of firms in the same industry) deters innovation while being close to firms from other industries stimulates innovation. This observation is in line with theories outlined in Arrow (1959) that competition between firms has a significant adverse effect on innovation. Given that patenting is a risky form of investment that generally requires significant financial and human resources, decisions that involve patentable innovation are made if justified by a benefit cost analysis. Firms in the same industry and in close proximity are often competing against each other in terms of intermediate goods, workforce, and customers. An increase in the number of firms in the same industry within an area increases competition within the area, which increases the level of uncertainty in future profitability. Once profitability is at risk, risk-adverse firms will not invest in riskier forms of assets (such as innovation), hence fewer patent activities. Locating close to innovative firms increases knowledge spillover, but this effect is only significant in a very small geographic range. The research findings in this study are important for a fast-developing country like China and suggest that to promote and increase innovative activities within Chinese industries, policymakers should foster policies that encourage a diversity of firms in areas that have a high density of firms such as science parks.

7. References

Acs, Z.J., Audretsch, D.B., and Feldman, M.P. (1994). R&D Spillover and the Recipient Firm Size. The Review of Economics and Statistics, 97(2): 336-340.

Aldieri, L. (2011). Technology and geographical proximity effects on knowledge spillovers: evidence from US patent citations. Economics of Innovation and New Technology, 20:597–607.

Arrow, K.J., (1959). Economic Welfare and the Allocation of Resources for Invention. Santa Monica, CA: RAND Corporation. Retrieve from: https://www.rand.org/pubs/papers/P1856.html

Audretsch, D. B. and Feldman, M. P. (1996a). R&D spillovers and geography of innovation and production. American Economic Review, 86:630–640 .

Audretsch, D.B., and Feldman, M.P. (1996b). Innovative Cluster and the Industry Life Cycle. Review of Industrial Organization, 11: 253-273.

Beaudry, C. and Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization verses urbanization debate. Research Policy, 38:318-337.

Bottazzi, L. and Peri, G. (2003). Innovation and spillovers in regions: Evidence from European patent data. European Economic Review, 47:687–710.

Carrión-Flores, C. E., Innes, R., & Sam, A. G. (2013). Do voluntary pollution reduction programs (VPRs) spur or deter environmental innovation? Evidence from 33/50. Journal of environmental economics and management, 66(3), 444-459.

Chan, K.W. and Zhang, L. (1999). The Hukou System and Rural-Urban Migration in China: Processes and Changes. The China Quarterly, 160: 818-855.

Chang, C. H., & Sam, A. G. (2015). Corporate environmentalism and environmental innovation. Journal of Environmental Management, 153, 84-92.

China Statistical Yearbook of Science and Technology (2019). China Data Insights. Retrieved from: http://cdi.cnki.net/YearData/Report/a621e297be23072

Chuang, Y. C., & Hsu, P. F. (2004). FDI, trade, and spillover efficiency: evidence from China's manufacturing sector. Applied Economics, 36(10), 1103-1115.

Eun, J.H., Lee, K., and Wu, G. (2006). Explaining the "University-run-enterprises" in China: A theoretical framework for university-industry relationship in developing countries and its applications to China. Research Policy, 35(9): 1329-1346.

Fact Book. (2011). Shanghai Stock Exchange. Retrieved from: http://english2019.sse.com.cn/indices/publications/factbook/c/factbook_us2011.pdf

Fan, C.C. (2002). The Elite, the Natives, and the Outsiders: Migration and Labor Market Segmentation in Urban China. Annals of the Association of American Geographers, 92(1): 103-124.

Fu, X. (2008). Foreign direct investment, absorptive capacity and regional innovation capabilities: Evidence from China. Oxford Development Studies, 36.

Fu, Y. and Gabriel, S.A. (2012). Labor migration, human capital agglomeration and regional development in China. Regional Science and Urban Economics, 42: 473-484.

Gao, T. (2004). Regional industrial growth: evidence from Chinese industries. Regional Science and Urban Economics, 34: 101-124.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., and Shleifer, A. (1992). Growth in Cities. Journal of Political Economy, 100(6): 1126-1152.

Hall, B. H. (2002). The financing of research and development. Oxford review of economic policy, 18(1), 35-51.

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.

Hericourt, J. and Poncet, S. (2009). FDI and credit constraints: Firm-level evidence from China. Economic Systems, 33: 1-21.

Hu, A. G. and Jefferson, G. H. (2009). A great wall of patents: What is behind China's recent patent explosion? Journal of Development Economics, 90:57–68.
Huang, C., and Wu, L. (2012). State-led Technological Development: A Case of China's Nanotechnology Development. World Development, 40(5): 970-982.
Huang, L., Lui, X., and Xu, L. (2010). Regional innovation and spillover effects of foreign direct investment in China: A threshold approach. Regional Studies,46.5:583–596.
Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: Evidence from firm's patents, profits, and market value. NBER Working Paper #1815.
Jaffe, A. B. (1989). Real effects of academic research. American Economic Review, 79:957–970.
Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers and evidenced by patent citations. The Quarterly Journal of Economics, 108:577–598.
Jefferson, G., Albert, G. Z., Xiaojing, G., & Xiaoyun, Y. (2003). Ownership, performance, and innovation in China's large-and medium-size industrial enterprise sector. China economic review, 14(1), 89-113.
Krugman, P. R. (1991). Geographic and Trade. MIT Press.
Li, X., Li, J., and Wu, X. (2019). University Spillovers, Spatial Distance, and Firm Innovation: Evidence at Chinese Listed Firms, Emerging Markets Finance and Trade, 1-16.
Lin, P., Liu, Z., and Zhang, Y. (2009). Do Chinese domestic firms benefit from FDI inflow? Evidence of horizontal and vertical spillovers. China Economic Review, 20: 677-691.
Liu, X. and Buck, T. (2007). Innovation performance and channels for inter-national technology spillovers: Evidence form Chinese high-tech industries. Research Policy, 36:355–366.
Liu, X., Lu, Y., Filatotchev, I., Buck, T., and Wright, M., (2010). Returnee entrepreneurs, knowledge spillovers and innovation in high-tech firms in emerging economies. Journal of International Business Studies, 41:1183-1197.
Liu, Z. (2002). Foreign direct investment and technology spillover: Evidence from China.Journal of Comparative Economics, 30:579 – 602.
Lucking, B., Bloom, N., and Reenen, J.V. (2018). Have R&D Spillovers Changed? National Bureau of Economic Research, NBER Working Paper Series #24622.
Moga, T.T. (2017). Patents in China. USPTO China IP Roadshow. Retrieved from: https://www.uspto.gov/sites/default/files/documents/11132017_Portland_RoadShow%28TomMoga%29.pdf
Ning, L, Wang, F., and Li, J., (2016). Urban innovation, regional externalities of foreign direct investment and industrial agglomeration: Evidence from Chinese cities. Research Policy, (45):830-843.
Rho, S. and Moon, I. (2014). Innovation and Spillovers in China: Spatial Econometric Approach. Seoul Journal of Economics, 27(2): 149-170.
Shang, Q., Poon, J.P.H., and Yue, Q. (2012). The role of regional knowledge spillovers on China's Innovation. China Economic Review, 23: 1164-1175.
Walcott, S.M. (2002). Chinese Industrial and Science Parks: bridging the Gap. The Professional Geographer, 54 (3): 349-364.
Wallsten, S. J. (2001). An empirical test of geographic knowledge spillovers using geographic information systems and firm level data. Regional Science and Urban Economics, 311:571–599.
Wang, C. C. and Wu A., (2016). Geographical FDI knowledge spillover and innovation ofindigenous firms in China. International Business Review, 25: 895-906.
Wang, Y., Ning, L., Li, J., and Prevezer, M. (2014). Foreign direct investment spillovers and the geography of innovation in Chinese regions: The role of regional industrial specialization and diversity. Regional Studies, 50:805–822.4
World Bank, (2003). World Bank Investment Survey, 2003. The World Bank. Washington, D.C. Available at: at: http://www.enterprisesurveys.org/rawdata.aspx.
Wu, X.G. and Treiman, D.J. (2007). Inequality and Equality under Chinese Socialism: The Hukou System and Intergenerational Occupational Mobility. American Journal of Sociology, 113(2): 415 – 445.
Xie, R., Yao, S., Han, F., & Fang, J. (2019). Land Finance, Producer Services Agglomeration, and Green Total Factor Productivity. International Regional Science Review, 0160017619836270.
Xu, X.P. and Sheng, Y. (2011). Productivity Spillovers from Foreign Direct Investment: Firm-Level Evidence from China. World Development, 40(1): 62-74.
Yang, C. H. and Lin, H. L. (2011). Openness, absorptive capacity, and regional innovation inChina. Environment and Planning, 44:333–355.
Yueh, L. (2009). Patent laws and innovation in china. International Review of Law and Economics, 29:304–313.
Zhang, Y. and Roelfsema, H. (2014). Globalization, Foreign Direct Investment, and Regional Innovation in China. Journal of International Commerce, Economics, and Policy. 5(3): 1-26.

Zhou, D., Li, S. and Tse, D.K. (2002). The impact of FDI on the productivity of domestic firms: the case of China. International Business Review, 11: 465-484.
Zhu, Shengjun, Chong Wang, and Canfei He. (2019). High-speed Rail Network and Changing Industrial Dynamics in Chinese Regions. International Regional Science Review (2019): 0160017619835908.

8. Appendix

Table A1: Industries Represented in our Dataset and their Codes

Sector	Code
Farming, Forestry, Animal Husbandry & Fishery	A
Mining	B
Manufacturing	C
Electricity, Heat, Gas and Water Production and Supply	D
Construction	E
Wholesale & Retail	F
Transportation, Storage and Post	G
Information Transmission, Software and IT services	I
Financial	J
Real Estate	K
Leasing and Business Services	L
Scientific Research and Technical Service	M
Water Conservancy, Environment, and Public Facilities Management	N
Education	P
Conglomerates	S