FDI Technology Spillovers and Spatial Diffusion in China*

Mi Lin\textsuperscript{a} and Yum K. Kwan\textsuperscript{a,b}

\textsuperscript{a} City University of Hong Kong
\textsuperscript{b} Wilfrid Laurier University

This version: 15 August 2012

Abstract: This paper investigates the geographical extent of FDI technology spillovers and diffusion in China. We employ spatial dynamic panel econometric techniques to detect TFP innovation clusters, to uncover the spatial extent of technology diffusion, and to quantify both the temporal and spatial dimension of FDI spillovers. Exploratory data analyses reveal that TFP innovations are positively correlated over space and this clustering pattern is getting stronger over time. Our empirical results show that FDI presence in a locality will generate negative and significant impact on the productivity performance of domestic private firms in the same location. Nevertheless these negative intra-regional spillovers are found to be locally bounded. Domestic private firms enjoy positive and significant FDI spillovers through inter-regional technology diffusion. Moreover, these inter-regional spillovers appear in spatial feedback loops among higher order neighboring regions. In the long run, the positive inter-regional spillovers outweigh the negative intra-regional spillovers, bestowing beneficiary total effect on domestic firms.

Key Words: FDI spillovers, spatial diffusion, spatial dynamic panel, Chinese economy.

JEL Classification: R12, F21, O33.

* We would like to thank Professors Cheng Hsiao, Chia-Hui Lu and Eden S.H. Yu for helpful discussions, comments and suggestions. All remaining errors are our own. This work is supported by the Research Center of International Economics of City University of Hong Kong (Project No. 7010009).

Correspondence: Mi Lin, The Research Center for International Economics, City University of Hong Kong, Kowloon, Hong Kong. Email address: linmixm@gmail.com
1. Introduction

Firms tend to agglomerate in specific areas so as to reduce transaction cost and exploit external economies (Marshall, 1920). The FDI location literature has documented the ensuing self-perpetuating growth or agglomeration pattern of multinational corporations (MNCs) over time (see, among others, Head et al., 1995; Cheng and Kwan, 1999a, 1999b; Blonigen et al., 2005; Lin and Kwan, 2011). The externalities arising from foreign direct investment (FDI) penetration also have long received great attentions from both economists and policy makers. Although the previous literature has provided some evidence of FDI spillovers at the firm and industry level in China (Lin et al., 2009; Abraham et al., 2010; Hale and Long, 2011; Xu and Sheng, 2012, among other earlier contributions), little is known about the extent to which the regional penetration of FDI affects the aggregate productivity of local private firms in spatial dimension. This paper studies FDI spatial spillovers using county-level data supplemented with precise GPS information of China. More specifically, this paper asks: Do domestic private firms benefit from FDI presence in their local and neighboring regions? What is the geographic extent of FDI spillovers? Do FDI spillovers attenuate with distance? If so, how rapid is the geographic attenuation pattern?

There is a vast literature on FDI spillovers. FDI may benefit domestic firms via channels like labor market turnover, new technology demonstration, local capital accumulation, competition in sales market, and ‘learning by watching’ opportunity for local firms (see, among others, MacDougall, 1960; Kokko, 1994; Fan, 2002; Blalock and Gertler, 2008). It has also been documented that the source of FDI, as well as the ability of local firms to absorb spillovers, matters for technology diffusion (Findlay, 1978; Kokko et al., 1996; Sjöholm, 1999; Javorcik, 2004; Abraham et al., 2010). FDI spillovers from MNCs to domestic firms can also be negative. A leading example is the ‘demand effect’ or ‘market stealing effect’ (Aitken and Harrison, 1999). In the short run indigenous firms may be constrained by high fixed cost which prevents them from reducing their total cost; therefore, foreign firms with cost advantages can steal market share from domestic firms via price competition. As a result, the shrinking demand will push up the unit cost of domestic firms and decrease their operation efficiency. Consequently, while the penetration of MNCs in the host country may bestow positive externalities on domestic firms, it could also introduce, at the same time, a negative demand effect, which drags down the productivity of local firms. The net impact from FDI presence on domestic firms depends on the magnitude of these two opposite externalities.

---

1 Marshall (1920) argues that firms can benefit from two types of external economies: 1) economies arising from ‘the use of specialized skill and machinery’ which depend on ‘the aggregate volume of production in the neighborhood’ and 2) economies “connected with the growth of knowledge and the progress of the arts” which tie to the ‘aggregate volume of production in the whole civilized world.’
Though the theoretical arguments are well established, the empirical literature so far provides mixed evidence in terms of the existence, the sign, and the magnitude of FDI spillovers.

It has been shown in the recent literature that results obtained for a sample of heterogeneous firms may reveal an incomplete and potentially misleading picture of the reality (Crespo, 2009). FDI spillovers may occur only among a sub-group of firms that have certain characteristics in common. More specifically, the diffusion and realization of spillovers from MNCs to domestic firms are not universal; instead, they can be affected by many factors drawn from both the economic and the geographical dimension (Nicolini and Resmini, 2011). On one hand, the absorptive capacity of a domestic firm would determine its propensity to engage in knowledge sharing as well as its likelihood to successfully assimilate foreign knowledge (Findlay, 1978; Glass and Saggi, 1998). On the other hand, geographical distance would determine the costs and then the attenuation pattern of technology diffusion, which may reduce the possibilities for indigenous firms that are distant from multinational enterprises to expropriate the spillovers. In this paper, we pursue this line of research by analyzing the role of FDI in the formation of TFP spatial autocorrelation process as well as the geographic extent of FDI spillovers.

Many analyses in previous literature traditionally assume that each region to be an isolated entity. The role of spatial dependence is neglected, even though it is an important force in the process of productivity growth (Rey and Montouri, 1999; Madriaga and Poncet, 2007) and ignoring spatial factors in empirical studies could result in serious misspecification when these factors are actually exist (Anselin, 2001; Abreu et al., 2005). Most previous studies on the impact of FDI on economic growth or the productivity upgrading of indigenous firms fail to take into account spatial interactions and, as a result, previous estimates and statistical inferences are questionable (Madriaga and Poncet, 2007; Corrado and Fingleton, 2012). In this paper, we empirically illustrate that model that fails to consider spatial factors may provide misleading results when these spatial factors do exist.

There is wisdom from theoretical literature motivating that regional total factor productivities (TFPs) might be spatially correlated. Ciccone and Hall (1996) illustrate that density of economic activity (defined as intensity of labor, human, and physical capital relative to physical space) would affect productivity in spatial dimension through externalities and increase returns. The reason is that, since the transportation of products from one stage to the next involves costs that increase with distance, the technology for the production of all goods within a particular geographical area will have increasing returns, i.e., the ratio of output to input will rise with density of economic activity, even
when technologies themselves are constant returns. Moreover, density of economic activity will also contribute to productivity through externalities that associated with the physical proximity of production. Consequently, the increase of spatial density of economic activity will result in aggregate increasing returns and the presence of spatially correlated TFPs as well. The same research also documented that, for US, more than half of the variance of labor productivity across states can be explained by differences in the spatial density of economic activity. Fingleton (2001) presents a hybrid growth model with both features from the new economic geography theory and the endogenous growth theory. Specifically, the model assumed that technical progress depends on peripherality, urbanization, the diffusion of innovations to regions with low technology levels, and on pure spatial externalities. Consequently, increasing returns to scale in the production of intermediate goods relevant to manufacturing output offsets diminishing returns associated with congestion effects. The model is supported by the empirical results based on the data of European Union. The estimated results reveal a significant increase in the level of positive spatial autocorrelation of productivity growth over time. Moreover, it is argued that the spatial polarization in productivity will persist, since barriers to trade between regions has been lowered over time which will enhance spatial externalities, as labor markets become wider and local economies more integrated resulting in stronger spillover across regional boundaries. In the data explanatory analysis of this paper, we will present in details the evidence of spatial autocorrelation for both level and growth rate of regional TFPs in China.

Many recent studies emphasize the role of labor market pooling in the process of spatial knowledge spillovers. Fallick et al. (2006) and Freedman (2008) illustrate that industry co-agglomeration facilitates labor mobility (moving among jobs). Ellison et al. (2010) further document that industries employing the same types of workers tend to co-agglomerate. Duranton and Puga (2004) explore the micro-foundations based on spatial externalities arising from sharing, matching and learning among individuals. Kloosterman (2008) and Ibrahim et al. (2009) both argue that industry agglomeration promote knowledge spillovers since it facilities individuals to share ideas and tacit knowledge. In line with these studies, we adopt regional employment share as proxy for FDI spatial knowledge spillovers in this paper.

While a common theme in the existing literature is that agglomeration would promote spatial spillovers, the direction of association between geographical distance and spillovers, however, is not as clear as one would expect. Backed by the argument that the exchange of tacit knowledge requires face-to-face contacts, Audretsch and Feldman (1996) and Gertler (2003) emphasize that knowledge sharing is highly sensitive to geographical distance and geographical proximity would
promote knowledge spillovers. Boschma and Frenken (2010), nevertheless, have proposed the so-called proximity paradox, i.e., though geographical proximity may be a crucial driver for economic agents to interact and exchange knowledge, too much proximity between these agents in other dimensions might harm their performance. For instance, if two firms in the same locality have large overlapping in term of their knowledge bases, this high cognitive proximity generally implies that two firms have very similar competences, which means that their engaging in knowledge exchange would lead to a serious risk of weakening their competitive advantage with respect to the local network partner. Broekel and Boschma (2012) hence argue that it is not that the quantity of contacts and intensity of knowledge exchanges matters. The type of knowledge exchanged and how the exchanged knowledge matches the existing knowledge base of the firms may matter more. Broekel et al. (2010) also document that geographical proximity may reduce the innovative performance of a firm if there exist a dominance of local linkages, i.e., a innovating firm excessively engage in inter-regional cooperation but at the same time lack sufficient intra-regional linkages. As for organizational proximity, profit organizations have an interest to keep their knowledge away from competitors, while non-profit organizations like universities have a public mission and, therefore, are more willing to exchange knowledge with others. Consequently, in the context of current paper, while geographical proximity increases the likelihood of learning and knowledge sharing between domestic private firms and MNCs, similarities in the market they serve (Aitken and Harrison, 1999), their knowledge bases, and their organizational proximity may also result in negative impact on the productivity performance of domestic private firms.

The rest of this paper is organized as follows. Section 2 describes the data and presents the results from exploratory spatial analyses. Section 3 presents a spatial dynamic panel model that incorporates the spatial features observed in the data. Section 4 discusses various econometric issues and presents empirical results. The final section concludes with a summary and suggestions for future research.

2. Data and Exploratory Analysis

2.1 Data

Data employed in this paper come from the annual census of above-size manufacturing firms conducted by the National Bureau of Statistics (NBS) of China from 1998 to 2007 (known as the Chinese Industrial Enterprises Database, NBS-CIE database henceforth). The database includes firm-level census data for state-owned firms and non-state-owned firms with sales revenue over 5 million RMB. There are several variables (including the Chinese standard location indicator,
province code, city code, county code, district code, as well as firms’ full address) can help us to identify the location of a firm. Of all these variables, province code, city code and county code are most complete and consistent over years. Measures specifying the distance between individual firms are not available. We hence define ‘region’ as a county in this paper. Consequently, all variables in this paper are aggregate county level data from an unbalanced panel data set with 1379 counties in 1998 and 2133 counties in 2007, respectively. The longitude and latitude data of China’s administration division at county level obtained from the GADM database of Global Administrative Areas functions as a supplement to the NBS-CIE database for spatial data exploratory and regression analysis.\(^2\)

The first step of our data analysis is to estimate the total factor productivity at firm and then at the county level for later spatial data exploratory analysis and regression use. More specifically, we first estimate the firm level TFP by following Levinsohn and Petrin (2003) approach. The county level TFP is the weighted average of firm level TFP in the same county with the weight being the firm’s value added share in the underlying county. Brandt et al. (2012) addresses data prepare and cleaning issues of NBS-CIE database thoroughly. We follow the data cleaning strategy suggested by their study. We also make use the industry concordances, deflators for all nominal variables provided by the same paper. Nevertheless, since we do not have 1993 annual enterprise survey and investment deflator from 1998 to 2007 mentioned in their paper, we are not be able to replicate and obtain the real capital stock as in Brandt et al. (2012). We use the sum of circulating funds and net value of fixed assets as proxy for capital input. The capital input data is deflated by the investment price index obtained from the price information obtained from various issues of China Statistical Yearbook. Our results show that this deviation will not affect the TFP estimation too much. The following table compares our estimates of the national aggregate TFP growth with those reported in Brandt et al. (2012). Our estimates are very close to their results. In the later analysis, our TFP are estimates based on value-added production function.

<table>
<thead>
<tr>
<th>Period</th>
<th>Brandt et al. (2012, JDE): Figure 4</th>
<th>Our results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value-added Function</td>
<td>Revenue Function</td>
</tr>
<tr>
<td>1998-2007</td>
<td>7.96%</td>
<td>2.85%</td>
</tr>
</tbody>
</table>

Notes: Firm level TFP is estimated by following Levinsohn and Petrin (2003). The national aggregate TFP is the weighted average of firms’ TFP for each year with the weight being the value added share of a firm in that particular year.

\(^2\) See Appendix 1 for details about the administrative division of China and other geographic information at county level.
Domestic private firms in this paper are firms that do not receive capital funds from foreign investors or from any level of China’s government. FDI in the NBS-CIE database include foreign firms from Hong Kong, Macao and Taiwan (HMT-type FDI henceforth) and foreign firms that are not from HMT areas (F-type FDI henceforth). Appendix 2 reports the information of firms’ ownership structures and their portions in each year in the database. More specifically, our definition of domestic private firms is corresponding to the sum of firms with ownership structures from column (1) to column (7). F-type FDI is the sum of pure F-type FDI and Sino-F Joint Ventures (JVs). HMT-type FDI is the sum of pure HMT-type FDI and Sino-HMT JVs.

2.2 Exploratory Analysis

In this section, we present exploratory analysis of our data. Our focus is to present and reveal the salient features of spatial autocorrelation for variable of interests, i.e., TFP level and TFP growth rate of China. By definition, spatial autocorrelation describes the coincidence of value similarity with locational similarity (Anselin, 2001). Positive spatial autocorrelation means high or low values of a variable tend to cluster together in space, and negative spatial autocorrelation indicates high (low) values are surrounded by low (high) values. As standard measures, both global and local Moran’s I statistic are commonly adopted in the literature to illustrate the strength and significance of spatial autocorrelation. Global Moran’s I statistic is defined as

\[
I_t = \frac{n}{S^2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_{i,t} - \mu_t)(x_{j,t} - \mu_t) \sum_{i=1}^{n} (x_{i,t} - \mu_t)^2
\]

where \(x_{i,t}\) is the variable of interest (TFP) for county \(i\) at time \(t\); \(\mu_t\) is the mean of variable \(x\) at year \(t\); \(w_{ij}\) is the element of spatial weights matrix \(W\) which will be formally defined in the next section. Notice that \(w_{ij}\) essentially functions as a weight to depict the relative similarity of two localities in terms of space. \(n\) is the number of counties. \(S^2\) is a scalar factor equal to the sum of all elements of spatial weights matrix \(W\). Similarly, local Moran’s I statistic is defined as

\[
I_{it} = \frac{(x_{i,t} - \mu_t) \sum_{j=1, j \neq i}^{n} w_{ij} (x_{j,t} - \mu_t)}{S^2} \quad \text{where} \quad S^2 = \frac{\sum_{j=1, j \neq i}^{n} (x_{j,t} - \mu_t)^2}{n-1} - \mu_t^2.
\]

This study does not attempt to address and evaluate the impact of FDI on the productivity of China’s state-owned enterprises. This issue may be investigated in future research.
Local Moran’s $I$ is also known as an example of Local Indicators of Spatial Association (LISA) in Anselin (1995). For both global and local Moran’s $I$, a positive value for $I$ statistic indicates that a county has neighboring counties with similarly high or low attribute values; this county is part of a cluster. A negative value for $I$ statistic indicates that a county has neighboring counties with dissimilar values; this county is an outlier. By comparing equations (1) and (2), it is straightforward to show that, for a row standardized weights matrix, the global Moran’s $I$ equals the mean of the local Moran’s $I$ statistics up to a scaling constant. Finally, both local and global Moran’s $I$ statistics require underlying variable is normally distributed. We employ normality test suggested by Shapiro and Francia (1972) and perform statistic test on both TFP level and growth rate. At the 5% significance level, the null hypothesis that the value of interest is normally distributed cannot be rejected.

Table 1 reports the global Moran’s $I$ statistics for aggregate county level ln(TFP) and TFP growth rate. As shown in the table, Moran’s $I$ statistics are significant and positive in all cases, implying the presence of positive spatial autocorrelation for both ln(TFP) and TFP growth rate. Notice that the statistics for domestic privates’ ln(TFP) increase significantly over time, indicating a enhancing process of spatial clustering in terms of TFP innovation for domestic private firms during the sample period.

### Table 2: Global Moran’s $I$ Statistics

<table>
<thead>
<tr>
<th></th>
<th>Moran’s $I$</th>
<th>Standard Deviation</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Firms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFP growth (1998-2007)</td>
<td>0.2194</td>
<td>0.0134</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ln(TFP) in 1998</td>
<td>0.2178</td>
<td>0.0131</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ln(TFP) in 2003</td>
<td>0.1693</td>
<td>0.0105</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ln(TFP) in 2007</td>
<td>0.2678</td>
<td>0.0106</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>Domestic Private Firms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFP growth (1998-2007)</td>
<td>0.1434</td>
<td>0.0167</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ln(TFP) in 1998</td>
<td>0.1030</td>
<td>0.0146</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ln(TFP) in 2003</td>
<td>0.1451</td>
<td>0.0109</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ln(TFP) in 2007</td>
<td>0.2303</td>
<td>0.0112</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Notes: All statistics are calculated based on row-standardized spatial weights matrix with 10 nearest neighbors. Statistics for TFP growth rate (1998-2007) are calculated based on the balanced panel.

Equation (1) essentially describes the correlation between spatially weighted (spatial lag) variable, $Wz$, and $z$ itself, where $z$ is the standardized variable of interest (TFP). Consequently, Moran’s $I$ statistic can also be illustrated by plotting $Wz$ against $z$ while the statistic is equivalent to the slop coefficient of the linear regression of $Wz$ on $z$. Figure 1 presents the Moran scatterplot of TFP growth rate for a balanced panel from 1998 to 2007. In each graph, the four quadrants in the plot group the observations into four types of spatial interaction: high values located next to high values (high-high cluster in upper right-hand corner), low values located next to low values (low-low...
cluster in lower left-hand corner), high values located next to low values (high-low outlier in lower right-hand corner), and low values located next to high values (low-high outlier in upper left-hand corner). Though there is no clear pattern of plots for TFP growth, Figure 1 still reveal clearly a positive association between $Wz$ and $z$.

For the Moran scatterplots for ln(TFP) in Figure 2, however, pattern is much apparent. Since variables are standardized, plots over time are comparable. It is clear that, over time, there is a tendency that most observations are located in the upper-right quadrants, corresponding to high-high values. Consequently, the data shows clearly that the spatial distribution of TFP level is becoming more clustered.

Figure 3 shows the counties with statistically significant ($p$-value < 0.05) values of the local Moran’s $I$ statistic for TFP growth rate. The color code on the map indicates the corresponding quadrant in the Moran scatterplot (Figure 1) to which the counties belong. The graphs show clearly that there are several spatial clusters in terms of TFP growth rate. For TFP growth of all firms, there are five major high-high clusters including 1) northwestern and southwestern China (around provinces of Shanxi, Sichuan, Guizhou, and Yunnan); 2) northeastern China (around provinces of Heilongjiang and Jilin); 3) some areas in Inner Mongolia; 4) northwestern China (mainly in provinces of Ningxia and Gansu), and 5) northern part of Xinjiang Province. The low-low cluster and high-low outliers are mainly in southern and eastern coastal regions. For TFP growth for domestic private firms, however, the high-high clusters are mainly located in 1) south central part of China (around provinces of Hubei, Hunan, and Sichuan); 2) provinces of Shandong and Shanxi; and 3) northeastern China (around provinces of Heilongjiang and Jilin). There is no apparent low-low cluster of TFP growth for domestic private firms.
Figure 1: Moran Scatterplot of TFP Growth Rate

Figure 2: Moran Scatterplot of ln(TFP) in 1998 and 2007
Figure 3: Local Indicator of Spatial Association Cluster Map of TFP Growth Rate

LISA Cluster Map of TFP Growth for All Firms (1998-2007)

LISA Cluster Map of TFP Growth for Domestic Private Firms (1998-2007)
Figure 4 presents a comparison of local Moran statistic for ln(TFP) of domestic private firms between 1998 and 2007. The graphs show significant change of clustering location during the sample period. In 1998, there are only several clusters covering limited regions. The high-high clusters are mainly in 1) the province of Yunnan; 2) around the provinces of Shanxi, Henan, and Hebei; and 3) some areas in Inner Mongolia. There are also high-low outliers or low-low clusters in 1) provinces of Guanxi and Guandong and 2) provinces of Heilongjiang and Jilin. In 2007, however, high-high clusters spread over almost central and central-northern parts of China while the high-low outliers and low-low clusters shift to southern parts of China. It is apparent that, for TFP level of domestic private firms, the locations of clusters spread to broader regions over time and the spatial clustering pattern become much salient in 2007.

Finally, the distribution of FDI in China also has strong spatial pattern. The regional presence of FDI is largely affected by the related policy in China. The Special Economic Zones (SEZs), which mainly located at coastal areas, were set up in China in the early 1980s to attract foreign capital by exempting MNCs from taxes and regulations. In view of the success of this experiment, similar schemes, such as Open Coastal Cities (OCCs), Open Coastal Areas, Economic and Technological Development Zones (ETDZs) and Hi-Tech Parks, were also set up to cover broader and inner regions in the later years. Figure 5 compares the FDI spatial density distribution (measured as fixed asset share of FDI in a specific county) between 1998 and 2007. As shown in the graphs, FDI presence in 1998 mainly clusters in costal and central regions of China. The graph for 2007 indicates that the clustering pattern getting stronger over time. While the FDI presence spreads broader and inner areas in China, the clusters remain in costal and central regions of China. Notice that the magnitudes of density also become larger over time, indicating strong FDI self-reinforcing pattern in spatial dimension.

To sum up, data exploratory analysis reveals salient spatial autocorrelation feature for both TFP growth and ln(TFP). There is strong tendency that ln(TFP) for domestic private firms are getting more clustered throughout the sample period. The data also reveal strong FDI clustering and self-reinforcing pattern in spatial dimension. In the next section, we further explore these results in a spatiotemporal model that incorporates both spatial interactions across regions and technology diffusion of FDI.
Figure 4: Local Indicator of Spatial Association Cluster Map of ln(TFP) for Domestic Private Firms

LISA Cluster Map of ln(TFP) for Domestic Private Firms in 1998

LISA Cluster Map of ln(TFP) for Domestic Private Firms in 2007
Figure 5: FDI Spatial Density Distribution at County Level

FDI Density (Fixed Asset Share) at County Level (1998)

FDI Density (Fixed Asset Share) at County Level (2007)
3. Spatiotemporal Model and Its Steady State Representation

3.1 Spatiotemporal Model

Following LeSage and Pace (2009), we generalize the well-known partial adjustment model by assuming that the variable of interest, $\ln(TFP)$ of domestic private firms, in a specific region is influenced by its own and other regions’ past period values. The additional assumption essentially incorporates both the time and space dependence on past decisions of neighboring economic agents in the model, since cross-sectional spatial dependence could arise from a diffusion process working over time rather than occurring simultaneously (LeSage and Pace, 2009), which is consistent with the pattern we observe in the previous section. We show that a spatial partial adjustment mechanism can thus result in a long-run equilibrium characterized by simultaneous spatial dependence and time-space interactions, which justifies the use of time-space panel model specification (spatiotemporal model) as a platform for empirical analysis in this paper.

As a starting point, equation (3) illustrates the equilibrium and its determinants of a spatial version of partial adjustment model, where $y^*_i$ is the equilibrium $\ln(TFP)$ of domestic private firms in county $i$ at time $t$; $X_i$ is a $n \times p$ matrix containing $p$ potential determinants, including proxies for FDI penetration; $\alpha$ is a constant term and $\iota_n$ is a $n \times 1$ vector with all elements are 1; $W$, which will be further defined shortly, is a $n \times n$ symmetric spatial weights matrix depicting the manner of spatial interactions between county $i$ and its neighbors.

$$ y^*_i = X_i\psi + WX_i\zeta + \alpha + \iota_n $$

(3)

Consequently, the equilibrium level of the dependent variable, $y^*_i$, depends on the own (in the same county) observations of explanatory variables ($X_i$) and neighboring (spatial lag) observations ($WX_i$), and an intercept. The parameter $\psi$ captures the own-region effect of explanatory variables and $\zeta$ captures the effects of explanatory variables at neighboring localities.

We postulate a spatial partial adjustment process in equation (4), where $\phi$ governs the degree of partial adjustment between realized previous value of dependent variable, $y_{i,t-1}$ and the equilibrium value $y^*_i$; $\iota_n$ is an $n \times 1$ vector of disturbances which follows $N(0, \sigma^2 I_n)$ distribution; $I_n$ is an $n \times n$ identity matrix; parameters $\theta$ and $\pi$ respectively measure the extent of temporal and spatial dependence depict by the $n \times n$ matrix $G$. 
Substituting equation (3) into (4) yields equation (5),

\[ y_{it} = \phi Gy_{i,t-1} + X_{it}(1-\phi)\psi + WX_{it}(1-\phi)\xi + \alpha(1-\phi)t_{it} + \epsilon_{it} \]  

where the realized value of dependent variable, \( y_{it} \), depends on temporal and time-space lag of the dependent variable \( G_{y_{i,t-1}} = \theta y_{i,t-1} + \pi W y_{i,t-1} \), spatial lags of explanatory variables \( WX_{it} \), and conventional terms in partial adjustment model \( X_{it} \). For the ease of notation, equation (5) can be further simplified as

\[ y_{it} = G_1 y_{i,t-1} + X_{it} \beta + WX_{it} \gamma + c + \epsilon_{it} \]  

where

\[ G_1 = [\phi \theta I_n + \phi \pi W] = [\rho I_n + \rho^w W], \]

\[ G_2 = [(1-\phi)\psi I_n + (1-\phi)\xi W] = [\beta I_n + \gamma W], \]

\[ c = \alpha(1-\phi)t_{it}. \]  

Analogy to equation (5), the temporal and spatial autoregressive process are now governed by parameters \( \rho^t \) and \( \rho^w \) respectively and the parameters \( \beta \) and \( \gamma \) respectively measures the own-region (intra-regional) and neighboring-region (inter-regional) effects of explanatory variables, \( X_{it} \).

### 3.2 Steady State Representation

Equations (6) and (7) describe a classic spatiotemporal model. It is more convenient to work with a steady state representation of this model, since, as will be shown shortly, this steady state representation would greatly facilitate our empirical analysis when we bring the model to data. To work out the steady state representation of the model, we assume that the explanatory variables \( X_{it} \) grow at a constant rate \( (\phi) \) in each period

\[ X_{it} = \phi^t X_{i0} \]  

Other stability conditions include

\[ \rho^t \in [0,1), \]
\[ \rho^w \in [0,1), \]
\[ \text{and} \ (\rho^t + \rho^w) < \kappa, \]

where \( \kappa \) is a small positive constant. These conditions will ensure that for sufficiently large value of \( t \), \( G_1 \) will take on small values. Consequently, we have
\[
\lim_{t \to \infty} G_t' y_{i0} \to 0, \\
\text{and } \lim_{t \to \infty} G_t' \phi^{-t} \to 0.
\] (10)

By recursive substitution, equation (6) can be restated as:
\[
y'_i = G_t' y_{i0} + \left[ I_n \phi' + G_t \phi^{-t} + \ldots + G_t^{i-1} \phi \right] X_{i0} (\beta + \gamma) + c + \bar{c}
\] (11)

where
\[
\bar{c} = G_t^{i-1} c + \ldots + G_t c + c;
\]
\[
\bar{c} = G_t^{i-1} \varepsilon_i + \ldots + G_t \varepsilon_{i,t-1} + \varepsilon_u.
\] (12)

Given the stability conditions in equations (8) to (10), for sufficiently large \( t \), taking the expectation of dependent variable \( (y'_i) \) in equation (11) yields the long-term equilibrium of the model shown in equation (13)
\[
E(y'_i) \approx \left[ I_n + G_t \phi^{-1} + \ldots + G_t^{i-1} \phi^{-i+1} \right] \left[ \beta I_n + \gamma W \right]
\]
\[
\approx \left[ I_n - G_t \phi^{-1} \right]^{-1} \left[ \beta I_n + \gamma W \right]
\]
\[
\approx \left[ I_n - \phi^{-1} (\beta' I_n + \rho^W W) \right]^{-1} \left[ \beta I_n + \gamma W \right]
\]
\[
= \left[ I_n - \rho^W W \right]^{-1} \left[ \beta I_n + \gamma W \right]
\] (13)

where
\[
\rho^* = \frac{\rho^W}{\phi - \rho' \beta}, \quad \beta^* = \left( \frac{\phi}{\phi - \rho'} \right) \beta, \quad \text{and } \gamma^* = \left( \frac{\phi}{\phi - \rho'} \right) \gamma.
\] (14)

Notice that, for a specific explanatory variable \( (X'_r) \), the matrix
\[
S_r[W] = \left[ I_n - \frac{\rho^W}{\phi - \rho'} W \right]^{-1} \left[ \frac{\phi}{\phi - \rho'} \right] (\beta I_n + \gamma W)
\]
\[
= \left[ I_n - \rho^W W \right]^{-1} \left( \beta' I_n + \gamma' W \right)
\] (15)

is called spatial multiplier. Consequently, the impact on the expected value of dependent variable given changes in the \( r \)th explanatory variable \( (X'_r) \) is a function of the multiplier matrix \( S_r[W] \) as shown in equation (16)
\[
E(y'_i) = \sum_{r=1}^{p} S_r[W] x'_r
\] (16)
One of the virtues of the spatial econometrics is the ability to incorporate the multi-regional interactions in the regression. To achieve this goal, as shown in equations (15) and (16), these multi-regional interactions enter the model as spatially weighted average of the regressor, where the weights in this paper are based on the distance \((d_{ij})\) between two counties \(i\) and \(j\). This generates the spatial weights matrix at time \(t\), \(W_{N_t}\), which is a \(N_t \times N_t\) symmetric matrix with elements \(w_{ij}\) defined as

\[
w_{ij} = \begin{cases} 
(d_{ij})^{-1} & \text{if } i \neq j \\
0 & \text{if } i = j 
\end{cases}
\]  

(17)

where \(N_t\) denotes the number of observations in year \(t\). This spatial weights matrix is named as inverse-distance spatial weights matrix. To ensure that the spatial weights matrix is nonsingular even in large sample, minmax-normalizations is employed as suggested by Kelejian and Prucha (2010). In a minmax-normalized matrix \(\tilde{W}_{N_t}\), the elements now become

\[
\tilde{w}_{ij} = \begin{cases} 
\frac{w_{ij}}{\min\{\max_i(r_i), \max_j(c_j)\}} & \text{if } i \neq j \\
0 & \text{if } i = j 
\end{cases}
\]  

(18)

where \(\max_i(r_i)\) is the largest row sum of \(W_{N_t}\) and \(\max_j(c_j)\) is the largest column sum of \(W_{N_t}\). Notice that normalizing by a scalar preserves symmetry and the basic model specification. For the entire panel, the min-max-normalized matrix \(\tilde{W}\) is an \(n \times n\) symmetric matrix where \(n = \sum_{t=1}^{T} N_t\) is the sum of the number of observations across all years. As a robustness check, we will also report empirical results based on an alternative spatial weights matrix, i.e., inverse-distance matrix with fast spatial decay where the elements of the spatial weights matrix in equation (18) is replaced by \((1/d_{ij})^2\).

3.3 Interpretation of Parameter Estimates: Summary Measures of Impacts and Spatial Partitioning of Impacts

The estimated parameter of spatiotemporal model provides wealthier and more complicated information than conventional non-spatial regression. With spatial interactions been explicitly

---

4 Due to the data availability, we have an unbalanced panel with time span \(T = 10\). Notice that missing variable shouldn’t be accounted (should be deleted) when constructing spatial weights matrix. Otherwise, we will assign wrong weights to spatially lag variables and lead to measurement error in spatial averaging terms. See Baltagi et al., (2008).

5 The major purpose of normalizing a spatial weights matrix is to provide boundaries so as to assure non-singularity. Kelejian and Prucha (2010) illustrate that row normalization, which uses different scalars across rows in spatial weights matrix, may lead to misspecification problem. They urge that, row normalization, unless can be justified by theoretical argument, should not be employed. Minmax-normalization, on the other hand, preserves the basic model specification.
incorporated, equation (16) implies that a change in a single observation (county) associated with any given explanatory variable will generate estimates measuring impact on the region itself (direct impact / intra-regional impact) and potentially impact on all other regions indirectly (indirect impact / inter-regional impact). This feature serves our purpose well as the main object of this paper is to detect and estimate the technology diffusion pattern in spatial dimension.

Equation (16) implies the following two derivatives given changes in the \( r \)th explanatory variable \( (X_{it}^r) \).

\[
\frac{\partial y_{it}}{\partial x_{ij}^r} = S_r[W]_{ij} \quad (19)
\]

\[
\frac{\partial y_{it}}{\partial x_{ij}^r} = S_r[W]_{ij} \quad (20)
\]

Equation (19) is named direct impact, which is the own derivative for the \( i \)th region in time \( t \), where \( S_r[W]_{ij} \) measures the impact on the dependent variable observation \( i \) from a change of \( x_{it}^r \) in county \( i \). Equation (20) is named total impact, which is the derivative of \( y_{it} \) with respect to \( x_{ij}^r \) for any \( i \) and \( j \). The difference between total effect and total direct effect is named total indirect effect. Taking average of these effects over all observations yields the following spatial summary measure of impacts

\[
Average \ Total \ Direct \ Impact \ (ATDE) = n^{-1}trace\left[ S_r (W) \right] \quad (21)
\]

\[
Average \ Total \ Impact \ (ATE) = n^{-1}t_n \left[ S_r (W) \right] t_n \quad (22)
\]

\[
Average \ Total \ Indirect \ Impact \ = ATE - ATDE \quad (23)
\]

Based on equations (19) and (20), it can be shown that the diagonal elements of the \( N_i \times N_i \) matrix \( S_r[W] \) contain the direct impacts, and off-diagonal elements contain indirect impacts. Notice that these impacts are calculated based on long-term equilibrium stated in equation (13), they should be interpreted as long-term equilibrium impacts.

The impacts described above include the effect of spatial feedback loops. For instance, a second order feedback effect means a change of observation \( x_{it}^r \) in county \( i \) affects observation in county \( j \), and county \( j \) also affects county \( i \). These feedback loops arise because county \( i \) is considered as a neighbor to its neighbors, so that impacts passing through neighboring counties will create a feedback impact on county \( i \) itself. This second order feedback effect is explicitly captured by the
terms in the third squared bracket of the last line in equation (15). The path of these feedback loops can be extended with the order of neighbors getting higher.

It is of interest to examine the profile of decaying magnitudes of impacts mentioned above when moving from lower-order neighbors to higher-order neighbors, as according to the first law of geography proposed by Tobler (1970): everything is related to everything else, but near things are more related than distant things. By investigating these profiles, we could explicitly reveal the extent to which the FDI spillovers spread over to neighboring regions as well as the rate of decay of these spillovers over space. The last line of equation (15) essentially provides guidance for these spatially partitioned effects, where the spatial summary measures of impacts are a function of \( S_r[W] \) which can be expanded as a combination of powers of the weights matrix using infinite series expansion of \((I_n - \rho W)^{-1}\). These powers correspond to the observations themselves (zero-order impacts with \( W^0 \) being the weight in the first squared bracket), immediate neighbors (first-order impacts with \( W^1 \) being the weight in the second squared bracket), neighbors of neighbors (second-order impacts with \( W^2 \) being the weight in the third squared bracket), and so on.\(^6\)

3.4 Computation Issues

In the empirical results, we report estimates of summary measures of direct, indirect and total impact as well as spatial partitioning of these impacts. To draw statistical inference on the significance of these impacts requires computation of standard errors of these impacts. This is not a straightforward task as the summary measures of effects are composed of different coefficients estimated according to complex mathematical formulas and the dispersion of these effects would depend on the dispersion of all estimated coefficients together. Following LeSage and Pace (2009) and Elhorst (2010), we use simulation method to obtain these standard errors by making use of the variance-covariance matrix of parameter estimates \((\hat{\rho}' , \hat{\rho}^W , \hat{\beta} , \hat{\gamma})\). More specifically, a random drawn from this variance-covariance matrix is

\[
\left[ \rho'_d , \rho^W_d , \beta_d , \gamma_d \right] = P^T \delta + \left[ \hat{\rho}' , \hat{\rho}^W , \hat{\beta} , \hat{\gamma} \right]^T
\]

where \(d\) denotes values obtained from draws; \(P\) is the upper-triangular Cholesky decomposition of \(Var(\rho' , \rho^W , \beta , \gamma)\); and \(\delta\) is a vector containing random values drawn from a standard normal

\(^6\) Notice that the main diagonal elements of the spatial weights matrix \(W\) are zeros, the main diagonal of higher order matrices \(W^m\) that arise in the infinite series expansion representation of the matrix inverse in equation (15), however, are non-zero. The main diagonal elements of \(W^*\), for instance, are nonzero to reflect the fact that region \(i\) is a second-order neighbor to \(i\) itself, that is a neighbor to its neighbor. This accounts for the feedback effects.
distribution. Each draw will result in one parameter combination for calculating impacts based on equations (21) to (23). We then obtain and report the mean value of impacts and corresponding standard errors by using 1,000 simulation draws.

Notice that these computations will involve taking inverse of the matrix \((I_n - \rho^*W)\) for every draw, which could be computationally inefficient especially when the dimension of the matrix is large. We follow the approach proposed by LeSage and Pace (2009) and use the following approximation instead of calculating the inverse of the matrix directly

\[
(I_n - \rho^*W)^{-1} = I_n + \rho^*W + (\rho^*)^2W^2 + (\rho^*)^3W^3 + \ldots + (\rho^*)^qW^q.
\] (25)

This would improve the computational efficiency since we only need to compute equation (25) once and could recall it easily when computing summary measures of impacts during simulation. Given this approximation, the formula in equations (21) to (23) could be simplified as shown in LeSage and Pace (2009). Denote

\[
T = \begin{bmatrix}
1 & 0 & n^{-1}tr(W^2) & n^{-1}tr(W^3) & \cdots & n^{-1}tr(W^q) \\
0 & n^{-1}tr(W^2) & n^{-1}tr(W^3) & n^{-1}tr(W^4) & \cdots & n^{-1}tr(W^{q+1})
\end{bmatrix}
\] (26)

\[
g = \begin{bmatrix}
1 & (\rho^*)^2 & \cdots & (\rho^*)^q
\end{bmatrix}
\] (27)

\[
G = \begin{cases}
g_i, & \text{diagonal elements} \\
0, & \text{off-diagonal elements}
\end{cases}
\] (28)

\[
P = \begin{bmatrix}
\beta & \gamma
\end{bmatrix}
\] (29)

then summary measures of impacts are now

\[
\text{Average Total Direct Impact (ATDE)} = PTGa
\] (30)

\[
\text{Average Total Impact (ATE)} = (\beta + \gamma)ga
\] (31)

\[
\text{Average Total Indirect Impact} = ATE - ATDE
\] (32)

where vector \(a\) is a \(1 \times (q+1)\) vector with all element are 1. As suggested by LeSage and Pace (2009), we set \(\max(q) = 100\). Notice that \(a\) could also be used as a control vector to obtain spatial partitioning of impacts. More specifically, by setting the \(q\)th element of \(a\) as 1 and all other elements as 0, equations (30) to (32) will yield \((q-1)\)-order summary measure of impacts. When all elements of \(a\) are 1, these equations will then provide accumulative (weighted average) impacts.

### 3.5 Variable Construction

The dependent variable \((y_{it})\) in this paper is weighted-average TFP of domestic private firms for each county. We first estimate firm level TFP industry by industry by following Levinsohn and
Petrin (2003) method. For each county \( i \) at time \( t \), \( y_{it} \) is the sum of weighted average of these firm level TFPs with the weight being the value added share of each firm in total value added for county \( i \) at time \( t \). FDI penetration is proxied by the employment share of foreign firms in the total employment of a county. We split foreign firms into two groups. They are foreign firms from Hong Kong, Macau, and Taiwan (HMT) and foreign from other countries or regions (F). More specifically, we have the following two proxies for FDI

\[
F_i^T = \frac{\text{Employment}_{it}^F}{\text{Employment}_{it}^T} \quad \text{and} \quad \text{HMT}_i^T = \frac{\text{Employment}_{it}^{\text{HMT}}}{\text{Employment}_{it}^T}
\]

where the superscript \( T \) denotes aggregate data for the whole county.\(^7\) We also construct two control variables to account for certain county heterogeneities. The variable \( SOE_{it} \) is a proxy for stated-owned enterprises’ presence in a county, which is defined as fixed assets share of stated-owned enterprises in a county, i.e.,

\[
(SOE \text{ Presence})_{it} = \frac{FA_{it}^{SOE}}{FA_{it}^T}
\]

where \( FA \) denotes fixed assets, and superscript \( T \) denotes aggregate data for the whole county. The second variable is a proxy for export intensity for a county, which is defined as

\[
(\text{Export Intensity})_{it} = \frac{\text{Export Value}_{it}}{\text{Gross Output}_{it}}
\]

Based on the variable constructed above, we have the following benchmark spatiotemporal regression function

\[
\ln(TFP)_{it} = \rho \ln(TFP)_{i,t-1} + \beta_1 (SOE \text{ Presence})_{it} + \beta_2 (\text{Export Intensity})_{it} + \\
\beta_3 F_i^T + \beta_4 \text{HMT}_i^T + \rho \tilde{W} \cdot \ln(TFP)_{i,t-1} + \gamma \tilde{W} \cdot (SOE \text{ Presence})_{it} + \\
\gamma \tilde{W} \cdot (\text{Export Intensity})_{it} + \nu_i + \lambda_t + v_{it}
\]

where \( \nu_i \) and \( \lambda_t \) are unobserved county-specific and time-specific effects, respectively. \( v_{it} \) is a pure random disturbance containing both time- and county- varying effects.

It is of interest to investigate the role of absorptive capacity of domestic private firms in the presence of potential spillovers. FDI is a combination of capital, technology and know-how from one country to another country. FDI per se can bring important benefits such as physical capital, advanced technology and improved managerial skills to the destination country. Nevertheless, it is

---

\(^7\) Employment share is commonly adopted as a proxy for FDI penetration in spillovers literature (see, among others, Aitken and Harrison, 1999; Abraham et al., 2010). We also try sales income share and fixed asset share as alternative proxies. Our results are robust to these alternative measurements. These results are available upon request.
argued that, these potential benefits do not automatically convert to spillovers. It is required that the
host country has sufficient capacities so as to facilitate the realization of spillovers. Although in the
literature there is strong evidence showing that there are potential FDI spillovers, there is also
ample evidence indicating that spillovers may not be automatic consequences of FDI penetration
countries should obtain a minimum level of absorptive capacity before exploiting the benefits from
FDI; otherwise, little can be expected from FDI.

To examine the association between absorptive capacity and FDI spillovers, we construct a dummy
We first calculate the R&D density of a firm, which is defined as the domestic private firm’s R&D
expenditures share in the total output. We then aggregate them to county level by using the
weighted average approach, where the weight is the value added share of each firm in total value
added for county i at time t. For a county, if its 3 years’ mean R&D intensity is higher than the
mean of all counties within the same time span, its dummy will take the value 1 and 0 otherwise.
We then construct interactive terms between these dummies and the proxies for FDI penetration to
examine the potential association between absorptive capacity and FDI spillovers.

4. Estimation Issues and Empirical Results

4.1 Estimation Issues

Kukenova and Monteiro (2009) propose to use the system-GMM estimator (Arellano and Bover,
1995; Blundell and Bond, 1998) to estimate a dynamic spatial panel model. Their research performs
Monte-Carlo investigation and compares the performance of system-GMM with various other
spatial dynamic panel estimators in terms of bias, root mean squared error and standard error
accuracy. In the scenario that accounts for endogeneity problem, their results are in favor of the
system-GMM estimator. Jacobs, Ligthart and Vrijburg (2009) also perform a Monte-Carlo study on
the same topic but their research allows for the presence of both spatial lag and spatial error in the
model. Estimators proposed by Kelejian and Robinson (1993) and Kapoor, Kelejian and Prucha
(2007) are also invited for the performance comparison. Results of Jacobs et al. (2009) confirm the
conclusion of Kukenova and Monteiro (2009), i.e., system-GMM out-performs other estimators.
Moreover, their Monte-Carlo evidence indicates that when system-GMM is adopted, differences in
bias, as well as root mean squared error, between spatial GMM estimates and corresponding GMM

8 These estimators include spatial MLE, spatial dynamic MLE (Elhorst, 2005), spatial dynamic QMLE (Yu et al., 2008),
LSDV, difference-GMM (Arellano and Bond, 1991), and system-GMM (Arellano and Bover, 1995; Blundell and
Bond, 1998).
estimates that ignore spatial correlation in error term are small. This research also documents that the combination of collapsing the instrument matrix and limiting the lag depth of the dynamic instruments substantially reduces the bias in estimating the spatial lag parameter, but hardly affects its root mean squared error. In view of these recent developments in econometric literature, all models reported in this section are estimated by using the spatial system-GMM estimator. The setup of moment conditions follow Kelejian and Prucha (1999), i.e., both spatially lagged dependent variable and independent variables are included in the instrument list on top of conventional instrument set for system-GMM suggested by (Arellano and Bover, 1995; Blundell and Bond, 1998).

4.2 Empirical Results

Table 3 reports the estimation results of both spatiotemporal model based on equation (36) and conventional dynamic panel model without spatial effects. Both time and spatial autocorrelation coefficients are positive and significant under different model specifications, suggesting fairly strong time and spatial self-reinforcing effects of total factor productivity for domestic private firms at county level. Estimated coefficients of proxies for own-regional (intra-regional) FDI presence are negative and significant and estimated coefficients of proxies for neighboring-regional (inter-regional) FDI presence are positive and significant across different regression models. Notice that, however, the absolute magnitude of the coefficients for intra-regional FDI presence proxies are lower in model without spatial effects, suggesting that conventional regressions ignoring spatial interactions may under estimate the negative direct (intra-regional) impact of FDI penetration.

<table>
<thead>
<tr>
<th>Dependent variable: ln(TFP)</th>
<th>No spatial effects</th>
<th>Inverse-distance matrix (1/d_{ij})</th>
<th>Inverse-distance matrix with fast spatial decay (1/d_{ij})^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time lag ln(TFP)</td>
<td>0.249***</td>
<td>0.131***</td>
<td>0.129***</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.039)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>SOE presence: FA</td>
<td>-4.022***</td>
<td>-7.371***</td>
<td>-7.283***</td>
</tr>
<tr>
<td></td>
<td>(0.856)</td>
<td>(0.765)</td>
<td>(0.893)</td>
</tr>
<tr>
<td>Export Intensity</td>
<td>-0.043</td>
<td>-1.226</td>
<td>-0.910</td>
</tr>
<tr>
<td></td>
<td>(0.795)</td>
<td>(0.994)</td>
<td>(1.044)</td>
</tr>
<tr>
<td>F presence: Employment</td>
<td>-5.938***</td>
<td>-9.352***</td>
<td>-8.753***</td>
</tr>
<tr>
<td></td>
<td>(1.671)</td>
<td>(1.795)</td>
<td>(1.807)</td>
</tr>
<tr>
<td>HMT presence: Employment</td>
<td>-4.163***</td>
<td>-10.336***</td>
<td>-10.406***</td>
</tr>
<tr>
<td></td>
<td>(1.406)</td>
<td>(1.625)</td>
<td>(1.604)</td>
</tr>
<tr>
<td>Space-time lagged of ln(TFP)</td>
<td>0.101*</td>
<td>0.140**</td>
<td>0.140**</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.061)</td>
<td></td>
</tr>
<tr>
<td>Spatially lagged Export Intensity</td>
<td>-18.996*</td>
<td>-37.349***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10.790)</td>
<td>(13.267)</td>
<td></td>
</tr>
<tr>
<td>Spatially lagged SOE presence: FA</td>
<td>-3.016</td>
<td>-4.017*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.979)</td>
<td>(2.335)</td>
<td></td>
</tr>
</tbody>
</table>
### Spatially lagged F presence: Employment

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hansen Statistic</td>
<td>7.133</td>
<td>26.084</td>
</tr>
<tr>
<td>Hansen Statistic P-value</td>
<td>0.211</td>
<td>0.098</td>
</tr>
<tr>
<td>D.O.F of Hansen Statistic</td>
<td>5.000</td>
<td>18.000</td>
</tr>
<tr>
<td>Number of Instruments</td>
<td>19.000</td>
<td>37.000</td>
</tr>
<tr>
<td>Arellano-Bond test for AR(1) in first differences</td>
<td>-11.195</td>
<td>-11.571</td>
</tr>
<tr>
<td>P-value for AR(1) Test</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Arellano-Bond test for AR(2) in first differences</td>
<td>0.973</td>
<td>-0.813</td>
</tr>
<tr>
<td>P-value for AR(2) Test</td>
<td>0.331</td>
<td>0.416</td>
</tr>
<tr>
<td>N</td>
<td>8642</td>
<td>8642</td>
</tr>
</tbody>
</table>

Notes: Results reported are two-step system-GMM estimates. Standard errors in parentheses. Windmeijer’s (2005) correction method for the two-step standard errors is employed. *p<0.10, ** p<0.05, *** p<0.01. Year dummies are included in all regressions. Collapsed instrument matrix technique is employed to reduce the instrument count.

To account for spatial feedback effects and draw inference from long-term equilibrium perspective, Table 4 reports statistics for direct (intra-regional), indirect (inter-regional), and total impacts based on 1,000 simulations. All summary measures of impacts are statistically significant. After accounting for the impact form spatial feedback loops, the intra-regional FDI spillovers are negative and the inter-regional FDI spillovers are positive; nevertheless, the inter-regional spillovers outweigh the intra-regional spillovers resulting in positive average total FDI spillovers in the long-term. In the second column of the Table 3, the estimated $\beta$ coefficient for F-type FDI presence is -9.352 while the corresponding average total direct impact in Table 4 is -10.322. The difference of these two magnitude, -0.97, captures the average feedback effects for intra-regional F-type FDI spillovers as a result of impacts passing through neighboring regions and back to the region itself. Notice that the dependent variable is (weighted average) ln(TFP) and FDI proxies are employment share; hence, the model is in semi-log form. Consequently, in the long-term, ceteris paribus, 1% increase of F-type FDI penetration in a county will on average associate with a 10.322% TFP drop of domestic private firms in the same locality, of which the spatial feedback effects account for around 9.4% \[ \left( \frac{0.97}{10.322} \right) \times 100\% \]. On the other hand, 1% increase of F-type FDI penetration in the neighboring regions of a typical county will associate with a 26.23% TFP increase of domestic private firms in this typical county, of which the impact of feedback loops account for 15.6%. Similarly, for HMT-type penetration, the average total direct (intra-regional) impact is -11.461% (feedback effects account for 9.8%) and average total indirect (inter-regional) impact is 28.934% (feedback effects account for 15.7%). The long-term average total impact for F-type and HMT-type FDI penetration are 15.907% and 17.473%, respectively. We obtain similar results for regression using fast decay spatial weights matrix.
### Table 4: Statistics of Impacts Based on Regression Results in Table 3

<table>
<thead>
<tr>
<th></th>
<th>Inverse-distance matrix (1/d&lt;sub&gt;ij&lt;/sub&gt;)</th>
<th>Inverse-distance matrix with fast spatial decay (1/d&lt;sub&gt;ij&lt;/sub&gt;)&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Summary Measures of Impacts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>HMT</td>
</tr>
<tr>
<td>Average total direct impact</td>
<td>-10.322*** (1.866)</td>
<td>-11.461*** (1.727)</td>
</tr>
<tr>
<td>Average total impact</td>
<td>15.907* (11.952)</td>
<td>17.473* (9.750)</td>
</tr>
</tbody>
</table>

|                      | Spatially Partitioned Impacts              |                                 |
|                      | F                                         | HMT                             | F                                         | HMT                             |
| Average total direct impact |  |  |  |  |
| W<sup>0</sup> | -10.281 | -11.409 | -9.608 | -11.467 |
| W<sup>1</sup> | 0.002 | 0.002 | 0.004 | 0.005 |
| W<sup>2</sup> | -0.000 | -0.000 | -0.000 | -0.000 |
| W<sup>3</sup> | 0.000 | 0.000 | 0.000 | 0.000 |
| W<sup>4</sup> | 0.000 | 0.000 | 0.000 | 0.000 |
| W<sup>5</sup> | 0.000 | 0.000 | 0.000 | 0.000 |
| Average total indirect impact |  |  |  |  |
| W<sup>0</sup> | 24.314 | 26.909 | 38.762 | 42.211 |
| W<sup>1</sup> | 1.071 | 1.236 | 3.079 | 3.389 |
| W<sup>2</sup> | 0.082 | 0.099 | 0.326 | 0.375 |
| W<sup>3</sup> | 0.000 | 0.001 | 0.004 | 0.005 |
| W<sup>4</sup> | 0.000 | 0.000 | 0.000 | 0.000 |
| W<sup>5</sup> | 0.000 | 0.000 | 0.000 | 0.000 |
| Average total impact |  |  |  |  |
| W<sup>0</sup> | 14.033 | 15.500 | 29.154 | 30.744 |
| W<sup>1</sup> | 1.073 | 1.238 | 3.083 | 3.394 |
| W<sup>2</sup> | 0.082 | 0.099 | 0.326 | 0.375 |
| W<sup>3</sup> | 0.000 | 0.001 | 0.004 | 0.005 |
| W<sup>4</sup> | 0.000 | 0.000 | 0.000 | 0.000 |
| W<sup>5</sup> | 0.000 | 0.000 | 0.000 | 0.000 |

Notes: All statistics reported are results from 1,000 simulations. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

Table 4 also reports the estimates for spatially partitioned impacts. The results indicate that intra-regional spillovers and inter-regional spillovers have very different spatial decay pattern. On average, the intra-regional spillovers become negligible even in first-order feedback loop (impact from immediate neighbors with W<sup>1</sup> being the weight). Notice that this first-order feedback is positive but very small in terms of magnitude, implying that the penetration of FDI in a specific county will affect its immediate neighbors, which in return will generate a positive but negligible feedback impact to the domestic private firms in this specific county. The inter-regional spillovers, however, are still significant even in second-order feedback loop (impact from neighbors of neighbors). When fast decaying spatial weights matrix is employed (i.e., more weights are located to closer neighbors) the inter-regional spillovers could even extend to the third-order feedback loop. These results suggest that the negative intra-regional FDI spillovers are bounded locally while the inter-regional FDI spillovers could extend to higher order neighbors. Specifically, FDI presence in a
county will generate significant negative spillovers to domestic private firms in the same locality. Moreover, these negative spillovers are contained in the underlying county and the magnitude of impact that extend to its neighbors through feedback loops is almost negligible, even for the underlying county’s immediate neighbors. Domestic private firms, however, mainly benefit from FDI penetration in their neighboring regions. These positive and significant FDI spillovers not only come from a county’s immediate neighbors but could also from its higher order neighbors. In the long run, the positive inter-regional spillovers outweigh the negative intra-regional spillovers in almost every spatially partitioned feedback loop, resulting overall total impact being positive and significant.

Table 5 and 6 report results using R&D intensity as proxy for absorptive capacity of domestic private firms. The estimates of the space-time lag parameter are significant, which confirm the importance of spatial interaction in the model. After accounting for impacts from spatial feedback loops, as shown in Table 6, when domestic private firms are low R&D intensive firms, we obtain the similar results as in benchmark regression. The intra-regional FDI spillovers are negative and significant while the inter-regional FDI spillovers are positive and significant and outweigh the intra-regional spillovers, which resulting in positive and significant FDI spillovers in overall level as the results from benchmark regression. When domestic private firms are high R&D intensive firms, intra-regional spillovers are still negative and significant for both F and HMT type FDI; nevertheless, F-type inter-regional FDI spillovers are still positive but not significant and the magnitude is smaller than intra-regional spillovers resulting in insignificant and negative average total impact. HMT-type FDI inter-regional FDI spillovers remain significant and positive, and outweigh the intra-regional spillovers as well; thus, the average total impact of HMT-type FDI are still positive and significant. Notice that high R&D intensive firms are usually firms with high productivity. Consequently, above results seems indicate that, when domestic private firms are low R&D intensive and the technology disparity between domestic private firms and FDI are large, domestic privates firms have more to learn and absorb the technology spillovers well through inter-regional spillovers. When domestic private firms are high R&D intensive, the asymmetric results for F-type and HMT-type spillovers may due to that when the gap of technology disparity between domestic privates FDI are small, the realization of spillovers may depend on other yet to be defined factors. One possible barrier may be that, compare to F-type FDI, HMT-type FDI may have knowledge base that match better with the knowledge base of high innovative intensive domestic private firms, while F-type FDI may have knowledge base that domestic private firms could not benefit from the spillovers directly, even though the gap of knowledge disparity is small.
### Table 5: Absorptive Capacity and FDI Spillovers - R&D Intensity

<table>
<thead>
<tr>
<th>Dependent variable: ln(TFP)</th>
<th>No spatial effects</th>
<th>Inverse-distance matrix (1/d_{ij})</th>
<th>Inverse-distance matrix with fast spatial decay (1/d_{ij})^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time lag ln(TFP)</td>
<td>0.274***</td>
<td>0.176***</td>
<td>0.187***</td>
</tr>
<tr>
<td>(0.044)</td>
<td>(0.039)</td>
<td>(0.040)</td>
<td></td>
</tr>
<tr>
<td>SOE presence: FA</td>
<td>-2.499***</td>
<td>-5.872***</td>
<td>-5.652***</td>
</tr>
<tr>
<td>(0.795)</td>
<td>(0.686)</td>
<td>(0.700)</td>
<td></td>
</tr>
<tr>
<td>Export Intensity</td>
<td>0.246</td>
<td>-0.022</td>
<td>-0.181</td>
</tr>
<tr>
<td>(1.059)</td>
<td>(0.782)</td>
<td>(0.745)</td>
<td></td>
</tr>
<tr>
<td>(2.279)</td>
<td>(1.179)</td>
<td>(1.113)</td>
<td></td>
</tr>
<tr>
<td>HMT presence: Employment</td>
<td>-4.863***</td>
<td>-7.194***</td>
<td>-7.402***</td>
</tr>
<tr>
<td>(1.849)</td>
<td>(1.143)</td>
<td>(1.074)</td>
<td></td>
</tr>
<tr>
<td>F presence * R&amp;D Dummy</td>
<td>-4.717*</td>
<td>3.662**</td>
<td>2.996**</td>
</tr>
<tr>
<td>(2.664)</td>
<td>(1.612)</td>
<td>(1.511)</td>
<td></td>
</tr>
<tr>
<td>HMT presence * R&amp;D Dummy</td>
<td>-2.508</td>
<td>-3.410*</td>
<td>-2.741</td>
</tr>
<tr>
<td>(2.488)</td>
<td>(1.891)</td>
<td>(1.723)</td>
<td></td>
</tr>
<tr>
<td>Space-time lagged of ln(TFP)</td>
<td>0.084**</td>
<td>0.065*</td>
<td></td>
</tr>
<tr>
<td>(0.040)</td>
<td>(0.038)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatially lagged Export Intensity</td>
<td>-30.122***</td>
<td>-28.148***</td>
<td></td>
</tr>
<tr>
<td>(7.795)</td>
<td>(7.320)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatially lagged SOE presence: FA</td>
<td>1.419</td>
<td>1.223</td>
<td></td>
</tr>
<tr>
<td>(1.749)</td>
<td>(1.746)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatially lagged F presence: Employment</td>
<td>34.227***</td>
<td>32.656***</td>
<td></td>
</tr>
<tr>
<td>(8.752)</td>
<td>(8.261)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatially lagged HMT presence: Employment</td>
<td>17.486**</td>
<td>18.301***</td>
<td></td>
</tr>
<tr>
<td>(7.016)</td>
<td>(6.672)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatially lagged F presence * R&amp;D Dummy</td>
<td>-31.196**</td>
<td>-30.272**</td>
<td></td>
</tr>
<tr>
<td>(13.905)</td>
<td>(13.086)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatially lagged HMT presence * R&amp;D Dummy</td>
<td>14.375</td>
<td>13.054</td>
<td></td>
</tr>
<tr>
<td>(12.168)</td>
<td>(11.373)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Hansen Statistic 9.687 54.633 63.874
Hansen Statistic P-value 0.207 0.154 0.219
D.O.F of Hansen Statistic 7.000 45.000 56.000
Number of Instruments 23.000 68.000 79.000
Arellano-Bond test for AR(1) in first differences -9.215 -10.862 -10.696
P-value for AR(1) Test 0.000 0.000 0.000
Arellano-Bond test for AR(2) in first differences 1.084 -0.330 -0.038
P-value for AR(2) Test 0.278 0.741 0.969
N 8661 8661 8661

Notes: Results reported are two-step system-GMM estimates. Standard errors in parentheses. Windmeijer’s (2005) correction method for the two-step standard errors is employed. *p<0.10, ** p<0.05, *** p<0.01. Year dummies are included in all regressions. Collapsed instrument matrix technique is employed to reduce the instrument count.

### Table 6: Statistics of Impacts Based on Regression Results in Table 5

<table>
<thead>
<tr>
<th></th>
<th>Low R&amp;D</th>
<th>High R&amp;D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summary Measures of Impacts</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average total direct impact</td>
<td>F</td>
<td>HMT</td>
</tr>
<tr>
<td>Low R&amp;D</td>
<td>-11.640***</td>
<td>-8.243***</td>
</tr>
<tr>
<td>High R&amp;D</td>
<td>(1.211)</td>
<td>(1.242)</td>
</tr>
<tr>
<td>Average total indirect impact</td>
<td>F</td>
<td>HMT</td>
</tr>
<tr>
<td>Low R&amp;D</td>
<td>41.019***</td>
<td>21.338***</td>
</tr>
<tr>
<td>High R&amp;D</td>
<td>(10.241)</td>
<td>(8.410)</td>
</tr>
<tr>
<td>Average total impact</td>
<td>F</td>
<td>HMT</td>
</tr>
<tr>
<td>Low R&amp;D</td>
<td>29.379***</td>
<td>13.095*</td>
</tr>
<tr>
<td>High R&amp;D</td>
<td>(9.770)</td>
<td>(7.677)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Low R&amp;D</th>
<th>High R&amp;D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summary Measures of Impacts</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average total direct impact</td>
<td>F</td>
<td>HMT</td>
</tr>
<tr>
<td>Low R&amp;D</td>
<td>-10.994***</td>
<td>-8.524***</td>
</tr>
<tr>
<td>High R&amp;D</td>
<td>(1.185)</td>
<td>(1.138)</td>
</tr>
<tr>
<td>Average total indirect impact</td>
<td>F</td>
<td>HMT</td>
</tr>
<tr>
<td>Low R&amp;D</td>
<td>38.334***</td>
<td>21.629***</td>
</tr>
<tr>
<td>High R&amp;D</td>
<td>(9.778)</td>
<td>(8.220)</td>
</tr>
<tr>
<td>Average total impact</td>
<td>F</td>
<td>HMT</td>
</tr>
<tr>
<td>Low R&amp;D</td>
<td>27.340***</td>
<td>13.105*</td>
</tr>
<tr>
<td>High R&amp;D</td>
<td>(9.256)</td>
<td>(7.551)</td>
</tr>
</tbody>
</table>
To summarize, we estimated a spatiotemporal model based on its long-run equilibrium representation. The space-time lag coefficient is positive and significant across different model specifications, indicating strong spatial autocorrelation of TFPs for domestic private firms, all other things being equal. This confirms the findings from exploratory analysis in Section 2. After accounting for spatial feedback loops, we obtain strong evidence showing that FDI presence in a county will impose negative and significant impact on productivity performance of domestic private firms in the same locality. Moreover, these negative intra-regional spillovers are bounded locally. These results may suggest that, when domestic private firms and MNCs are in the same locality, their proximities in other dimensions (i.e., the common market they serve, the same local partner, the overlapping knowledge based, and the same profit-oriented nature) may create negative impact that mitigate or even outweigh the benefit from geographical proximity. Domestic private firms, however, receive positive and significant spillovers form FDI through inter-regional knowledge diffusion. These inter-regional spillovers, different form intra-regional spillovers, can extend to higher order neighboring counties rather than bounded within a county. We show that models that fail to account spatial interactions may underestimate the negative intra-regional spillovers form FDI, which highlight the importance of incorporating spatial factors in the empirical studies when spatial features appear in the data. We also found that high R&D intensive firms may face barrier to absorb benefit from F-type FDI penetration.

5. Conclusions

In this paper, we investigate the geographic extent of FDI spillovers to domestic private firms in China. Previous literature has argued that the diffusion and realization of FDI spillovers are not automatic; instead, they can be affected by factors drawn from both the economic and the geographical dimension. We show that well-developed techniques in spatial econometrics can be
employed to detect TFP clusters and to present them visually, to reveal the spatial extend of technology diffusion, and to estimate empirical models that incorporate spatial interactions explicitly. By making use of the geographic information at county level, our data exploratory analysis reveals strong spatial autocorrelation for TFP growth and level in China. We then further explore these findings by present a spatiotemporal model incorporating both the spatial interaction for TFPs of domestic private firms and the FDI spillovers. We show that this spatiotemporal model can be justified by generalizing the well-known partial adjustment model by assuming that the variable of interest, ln(TFP) of domestic private firms, in a specific region is influenced by its own and other regions’ past period values. Consequently, a spatial partial adjustment mechanism can result in a long-run equilibrium characterized by simultaneous spatial dependence and time-space interactions. We work out the long-run equilibrium representation of this model and use it as the benchmark model in regression analysis.

Our empirical results confirm the finding from exploratory analysis. The time-space lag coefficient is positive and significant across different model specifications, indicating high or low TFPs tend to cluster together over time. Our estimations also reveal the geographical extend of FDI spillovers in terms of the sign, the magnitude, and the geographic attenuation pattern. Given all other things being equal, FDP penetration in one county will generate significant negative spillovers to the productivity performance of the domestic private firms in the same locality. These negative intra-regional FDI spillovers, as shown by the data, are bounded within the underlying county as the feedback effects from higher order counties are found to be negligible. Domestic private firms, however, obtain positive and significant FDI spillovers through inter-regional technology diffusion. Moreover, these inter-regional spillovers appear in spatial feedback loops among higher order neighboring counties. In the long run, the positive inter-regional spillovers outweigh the negative intra-regional spillovers, resulting in positive total effect. We also found that high R&D intensive firms may face barrier to absorb benefit from F-type FDI penetration and the sources of barrier may come from the proximities in other dimensions. For future research it would be interesting to explore explicitly the role of proximities in other dimensions that may also matter for FDI spillovers to local firms.
Appendix 1: Administrative division and geographic information of China at county level

About the administrative division, the People’s Republic of China (PRC) administers 34 province-level divisions, including 23 provinces, 5 autonomous regions (Guangxi, Inner Mongolia, Ningxia, Tibet, and Xinjiang), 4 municipalities (Beijing, Chongqing, Shanghai, and Tianjin), and 2 special administrative regions (Hong Kong and Macau). In China’s administrative division, provinces are theoretically subservient to the PRC central government, but in practice provincial officials have large discretion with regard to economic policy. The subsidiary cities, regions, counties, and villages are controlled directly by the upper divisions, while the upper leaders take the responsibility of the activities of the subsidiary. Under the Province level, as of 2009, there are Prefecture level, County level, Township level, and Village level making up by 333 prefectures, 40,859 township-level divisions, and cities, and 704,386 villages and committees. County-level divisions are the third level of local government, coming under both the province level and the prefecture level, including counties, autonomous counties, banners, autonomous banners, county-level cities and districts. There are 1,464 counties in Mainland China out of a total of 2,858 county-level divisions. The following table summarized the basic information for the first three level administrative divisions of China.

Table A1: Administrative Division of People’s Republic of China as of 2009

<table>
<thead>
<tr>
<th>Level</th>
<th>Name</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Provinces (省 shěng) (22)</td>
<td>Provinces (省 shěng) (22)</td>
</tr>
<tr>
<td></td>
<td>Claimed province (1)</td>
<td>Claimed province (1)</td>
</tr>
<tr>
<td></td>
<td>Autonomous regions (自治区 zìzhìqū) (5)</td>
<td>Autonomous regions (自治区 zìzhìqū) (5)</td>
</tr>
<tr>
<td></td>
<td>Municipalities (直辖市 zhíxiáshì) (4)</td>
<td>Municipalities (直辖市 zhíxiáshì) (4)</td>
</tr>
<tr>
<td></td>
<td>Special administrative regions (特别行政区 tèbié xíngzhèngqū) (2)</td>
<td>Special administrative regions (特别行政区 tèbié xíngzhèngqū) (2)</td>
</tr>
<tr>
<td>2</td>
<td>Prefectures (地区 dìqū) (17)</td>
<td>Prefectures (地区 dìqū) (17)</td>
</tr>
<tr>
<td></td>
<td>Prefecture-level cities (地级市 dìjíshì) (283)</td>
<td>Prefecture-level cities (地级市 dìjíshì) (283)</td>
</tr>
<tr>
<td></td>
<td>Autonomous prefectures (自治州 zìzhìzhōu) (30)</td>
<td>Autonomous prefectures (自治州 zìzhìzhōu) (30)</td>
</tr>
<tr>
<td></td>
<td>Leagues (盟 méng) (3)</td>
<td>Leagues (盟 méng) (3)</td>
</tr>
<tr>
<td>3</td>
<td>Counties (县 xiàn) (1,464)</td>
<td>Counties (县 xiàn) (1,464)</td>
</tr>
<tr>
<td></td>
<td>Districts (市辖区 shìxiáqū) (855)</td>
<td>Districts (市辖区 shìxiáqū) (855)</td>
</tr>
<tr>
<td></td>
<td>County-level cities (县级市 xiànjíshì) (367)</td>
<td>County-level cities (县级市 xiànjíshì) (367)</td>
</tr>
<tr>
<td></td>
<td>Autonomous counties (自治县 zìzhìxiàn) (117)</td>
<td>Autonomous counties (自治县 zìzhìxiàn) (117)</td>
</tr>
<tr>
<td></td>
<td>Banners (旗 qí) (49)</td>
<td>Banners (旗 qí) (49)</td>
</tr>
<tr>
<td></td>
<td>Autonomous banners (自治旗 zìzhìqí) (3)</td>
<td>Autonomous banners (自治旗 zìzhìqí) (3)</td>
</tr>
<tr>
<td></td>
<td>Special districts (特区 tèqū) (3)</td>
<td>Special districts (特区 tèqū) (3)</td>
</tr>
<tr>
<td></td>
<td>Forestry areas (林区 línqū) (1)</td>
<td>Forestry areas (林区 línqū) (1)</td>
</tr>
</tbody>
</table>


The geographic information, including longitude and latitude, of China’s administration divisions at county level are obtained from the GADM database of Global Administrative Areas, which is maintained by the International Rice Research Institute and the Museum of Vertebrate Zoology at
the University of California, Berkeley. GADM is a spatial database of the location of the world’s administrative areas (or administrative boundaries) and describes where these administrative areas are (the spatial features), and for each area it provides some attributes, such as the name, geography area, longitude and latitude, and shape.\(^9\) As for China, the latest data was updated in 2009 with information for 2,410 county-level divisions were incorporated in the database.

We merge GADM database and NBS-CIE database by using the county code and the name of county in these two databases. For each county in China, there is a unique county code assigned by China’s administrative authority. GADM follows this system only that it does not distinguish districts directly governed by a city (i.e., districts within urban area and directly under the jurisdiction of municipal government of a city). As an example, here we illustrate a case based on the city of Beijing. Beijing is one of the four municipalities China, and is further divided into 18 county-level divisions (16 districts and 2 counties in 2009). Among these 18 county-level divisions, as shown in the figure below, 8 districts are within the urban area of the city of Beijing (Dongcheng, Xicheng, Chongwen, Xuanwu, Shijingshan, Haidian, Chaoyang, and Fengtai). Consequently, these 8 divisions are treated as a same locality and share a same set of longitude and latitude in the GADM database. After the merge, we obtain an unbalanced panel with 1,379 regions in 1998 and 2,133 regions in 2007.

**Figure A1: Administrative divisions of the city of Beijing**

The following table compares the average area of county-level divisions in our dataset with some other regions around the world. The average area of county-level divisions in China is about half of the average area of Japanese Prefectures and more than double of the average area of ceremonial counties of England.

Table A2: Some statistics for the area of administrative division around the world

<table>
<thead>
<tr>
<th>Division</th>
<th>Area (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average area of county-level divisions in China</td>
<td>3,949</td>
</tr>
<tr>
<td>Average area of Japanese Prefectures</td>
<td>7,959</td>
</tr>
<tr>
<td>Average area of second level divisions for Taiwan</td>
<td>1,559</td>
</tr>
<tr>
<td>Average area of ceremonial counties of England</td>
<td>2,714</td>
</tr>
<tr>
<td>Median land area of U.S. counties</td>
<td>1,610</td>
</tr>
</tbody>
</table>

Appendix 2: Information of Ownership Structure and Their Corresponding Portions in NBS-CIE Database

<table>
<thead>
<tr>
<th>Year</th>
<th>No. of Firms</th>
<th>Pure Privates</th>
<th>Other Domestic Firms</th>
<th>Collective Enterprise</th>
<th>Joint-Stock Enterprise</th>
<th>Associated Economics</th>
<th>Limited Liability Company</th>
<th>Corporation Limited Enterprises</th>
<th>All Domestic Privates</th>
<th>Pure SOEs</th>
<th>SOEs-Domestic JVs</th>
<th>SOEs</th>
<th>Pure F-type FDI</th>
<th>Joint Ventures between F and HMT</th>
<th>FDI</th>
<th>Sino-F-type JVs</th>
<th>Sino-HMT JVs</th>
<th>Other Sino-Foreign JVs</th>
<th>Sino-Foreign JVs</th>
<th>Sino-Foreign JVs</th>
<th>Undefined</th>
<th>Total: (8) + (11) + (15) + (19) + (20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>102830</td>
<td>11.25</td>
<td>3.36</td>
<td>30.44</td>
<td>1.54</td>
<td>2.49</td>
<td>1.18</td>
<td>34.70</td>
<td>16.48</td>
<td>5.87</td>
<td>22.35</td>
<td>(9)</td>
<td>2.84</td>
<td>3.88</td>
<td>0.03</td>
<td>6.75</td>
<td>6.43</td>
<td>6.76</td>
<td>0.27</td>
<td>13.46</td>
<td>2.74</td>
<td>100</td>
</tr>
<tr>
<td>1999</td>
<td>111715</td>
<td>13.17</td>
<td>4.53</td>
<td>27.44</td>
<td>4.48</td>
<td>1.43</td>
<td>3.17</td>
<td>13.9 55.60</td>
<td>14.80</td>
<td>5.96</td>
<td>20.76</td>
<td>(10)</td>
<td>3.13</td>
<td>4.97</td>
<td>0.04</td>
<td>8.14</td>
<td>5.89</td>
<td>6.60</td>
<td>0.25</td>
<td>12.75</td>
<td>2.76</td>
<td>100</td>
</tr>
<tr>
<td>2000</td>
<td>110090</td>
<td>19.10</td>
<td>5.80</td>
<td>22.25</td>
<td>4.08</td>
<td>1.19</td>
<td>4.27</td>
<td>1.47 58.17</td>
<td>11.99</td>
<td>5.34</td>
<td>17.33</td>
<td>(11)</td>
<td>3.41</td>
<td>5.50</td>
<td>0.05</td>
<td>8.96</td>
<td>5.77</td>
<td>6.84</td>
<td>0.23</td>
<td>12.85</td>
<td>2.69</td>
<td>100</td>
</tr>
<tr>
<td>2001</td>
<td>125235</td>
<td>26.56</td>
<td>7.86</td>
<td>16.84</td>
<td>3.66</td>
<td>0.95</td>
<td>5.57</td>
<td>1.52 62.96</td>
<td>8.97</td>
<td>4.32</td>
<td>13.29</td>
<td>(12)</td>
<td>3.83</td>
<td>5.82</td>
<td>0.05</td>
<td>9.70</td>
<td>5.44</td>
<td>6.19</td>
<td>0.22</td>
<td>11.85</td>
<td>2.21</td>
<td>100</td>
</tr>
<tr>
<td>2002</td>
<td>136163</td>
<td>32.63</td>
<td>9.12</td>
<td>13.26</td>
<td>2.84</td>
<td>0.71</td>
<td>5.98</td>
<td>1.52 66.06</td>
<td>7.26</td>
<td>3.65</td>
<td>10.90</td>
<td>(13)</td>
<td>4.28</td>
<td>5.69</td>
<td>0.07</td>
<td>10.04</td>
<td>5.34</td>
<td>5.36</td>
<td>0.20</td>
<td>10.89</td>
<td>2.10</td>
<td>100</td>
</tr>
<tr>
<td>2003</td>
<td>153585</td>
<td>37.94</td>
<td>10.95</td>
<td>9.38</td>
<td>2.22</td>
<td>0.52</td>
<td>6.46</td>
<td>1.47 68.95</td>
<td>5.30</td>
<td>2.70</td>
<td>8.00</td>
<td>(14)</td>
<td>4.72</td>
<td>6.35</td>
<td>0.05</td>
<td>11.12</td>
<td>5.01</td>
<td>4.89</td>
<td>0.16</td>
<td>10.06</td>
<td>1.88</td>
<td>100</td>
</tr>
<tr>
<td>2004</td>
<td>219335</td>
<td>44.28</td>
<td>11.83</td>
<td>5.59</td>
<td>1.56</td>
<td>0.31</td>
<td>7.70</td>
<td>1.21 72.47</td>
<td>3.21</td>
<td>1.76</td>
<td>4.97</td>
<td>(15)</td>
<td>5.72</td>
<td>6.58</td>
<td>0.05</td>
<td>12.36</td>
<td>4.79</td>
<td>4.11</td>
<td>0.13</td>
<td>9.03</td>
<td>1.18</td>
<td>100</td>
</tr>
<tr>
<td>2005</td>
<td>218401</td>
<td>44.93</td>
<td>13.76</td>
<td>4.47</td>
<td>1.28</td>
<td>0.26</td>
<td>7.44</td>
<td>1.19 73.34</td>
<td>2.44</td>
<td>1.37</td>
<td>3.81</td>
<td>(16)</td>
<td>5.97</td>
<td>6.88</td>
<td>0.05</td>
<td>12.91</td>
<td>4.41</td>
<td>3.68</td>
<td>0.10</td>
<td>8.19</td>
<td>1.76</td>
<td>100</td>
</tr>
<tr>
<td>2006</td>
<td>239352</td>
<td>47.68</td>
<td>14.12</td>
<td>3.75</td>
<td>1.36</td>
<td>0.21</td>
<td>7.08</td>
<td>1.01 75.21</td>
<td>1.89</td>
<td>1.09</td>
<td>2.98</td>
<td>(17)</td>
<td>5.89</td>
<td>6.70</td>
<td>0.05</td>
<td>12.64</td>
<td>4.09</td>
<td>3.31</td>
<td>0.09</td>
<td>7.49</td>
<td>1.69</td>
<td>100</td>
</tr>
<tr>
<td>2007</td>
<td>266311</td>
<td>48.43</td>
<td>16.27</td>
<td>2.96</td>
<td>1.14</td>
<td>0.17</td>
<td>6.90</td>
<td>0.92 76.78</td>
<td>1.42</td>
<td>0.86</td>
<td>2.28</td>
<td>(18)</td>
<td>5.67</td>
<td>2.91</td>
<td>0.09</td>
<td>6.67</td>
<td>3.67</td>
<td>2.91</td>
<td>0.08</td>
<td>6.67</td>
<td>1.75</td>
<td>100</td>
</tr>
</tbody>
</table>

Notes: Total record of 503079 firms with 1683017 observations. All statistics reported are results after data cleaning. A firm’s ownership structure is determined by its source and structure of Paid in Capital and their registered type. ‘Pure’ means the Paid in Capital is 100% from the corresponding source; for instance, ‘Pure Private’ means all the Paid in Capital of these firms are from privates.
References


