Agglomeration economies in Vietnam : a firm-level analysis

INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 636

Agglomeration economies in Vietnam: A firm-level analysis

Toshitaka GOKAN*, Ikuo KUROIWA** and Kentaro NAKAJIMA†

March 2017

Abstract

This paper examines the effects of agglomeration economies on firm-level productivity in Vietnam. By using Vietnamese firm-level data and the cluster detection method proposed by Mori and Smith (2013), we estimate the agglomeration economies for firm-level productivity. Specifically, we consider the different effects of agglomeration economies for localization and urbanization, as well as across types of firms; state-owned, private, and foreign-owned firms. Furthermore, we decompose the agglomeration economies into the three sources of the effect; inter-industry transaction relationships, knowledge spillovers, and labor pooling. We find the following results. First, localization economies actually improve firm-level productivity in Vietnam, with firms in the clustered areas having higher productivities. However, the localization economies do not improve the productivity of the state-owned firms. Second, urbanization economies improve productivity only for foreign-owned firms. State-owned and private firms do not benefit from urbanization economies. From the decomposition of agglomeration economies, we find that agglomeration economies formed through transactions work only for private firms. On the other hand, agglomeration economies formed through knowledge spillovers and labor pooling work for foreign-owned firms.

Keywords: Agglomeration Economies; Economic Geography

JEL classification: R12

* Researcher, Economic Geography Group, Area Studies Center, IDE (toshitaka_gokan@ide.go.jp)

** Executive Senior Research Fellow, Bangkok research center, IDE

† Associate Professor, Graduate school of Economics, Tohoku University

1. Introduction

Since Marshall's (1920) seminal study, the importance of agglomeration economies in improving productivity has been widely recognized, and many empirical papers have investigated the agglomeration economies found in many different countries. This research has, however, focused mainly on agglomeration economies in market-based countries like the US (e.g., Ciccone and Hall, 1996; Henderson, 2003), UK (e.g., Ciccone, 2002), with research on transition economies more rare. Recently, empirical analysis of agglomeration economies in the Chinese economy is increasing (e.g., Lin, Li, and Yang, 2011; Drucker and Feser, 2012). However, there are still few studies focusing on other transition countries.

Vietnam has experienced a transition from a centrally planned to a market-based economy since 1986, when the nation adopted a new economic reform policy called Doi Moi. Since the reform started, export-oriented labor-intensive manufacturing sectors, such as apparel and footwear, have developed rapidly thanks to massive inflows of foreign direct investment. Vietnam's economic growth was further boosted by trade liberalization, especially after Vietnam signed a bi-lateral trade agreement with the USA in 2000, followed by accession to the WTO in 2007. Vietnam's per capita GDP, once as low as 100 USD at the beginning of the reforms, exceeded 2,000 USD in 2014. Vietnam's economy had been affected by the global economic depression for some years since 2008, but recovered to around the same level as the pre-depression period within a rather short time. GDP growth rate in 2015 was 6.7%.

In spite of the large presence of the Vietnamese economy in transition economies, studies of spatial agglomeration in Vietnam are still rare. As an exception, Ercole (2013) investigates agglomeration in Vietnam, finding that a few regions lead the country's rapid economic growth, with economic activity highly concentrated in Ho Chi Minh City. Further, low-tech industries are more agglomerated than the mid-high- and high-tech industries. As another exception, Howard, Newman, and Tarp (2012) identify the determinants of agglomeration in Vietnam by applying the approach of Ellison, Glaeser, and Kerr (2010). They cannot find determinants of agglomeration robustly, with the identified determinants varying with the choice of the measure of agglomeration.

Ercole's (2013) result that economic activities are agglomerated only in low-tech industries, but the mid-high- and high-tech industries implies that agglomeration economies would not be well formed in Vietnam, especially in mid-high- and high-tech industries. Further, Howard, Newman, and Tarp's (2012) finding that the determinants of agglomeration are not identified robustly may come from the weak agglomeration effects in Vietnam. However, most of the studies on agglomeration in Vietnam focus only on the location patterns of economic activities with the regionally aggregated data, and there is no paper that examines effects of agglomeration on productivity.

Against this backdrop, this paper examines the effects of agglomeration economies on firm-level productivity in Vietnam. By using firm-level data, we can estimate firm-level TFP. We then identify the industrial clusters as a convex combination of contiguous districts as a proxy for agglomeration using the procedure proposed by Mori and Smith (2013). Then, by comparing TFP between cluster and non-cluster, we estimate the agglomeration effects on firm-level productivity. We especially consider the different effects of the agglomeration economies in localization and urbanization. Furthermore, a special characteristic of a transition economy like Vietnam is the large presence of state-owned and foreign-owned firms. This paper also investigates the difference in the strength of agglomeration economies across firm characteristics.

Furthermore, we decompose the channels of agglomeration economies working in the cluster. As Marshall (1920) pointed out, there are three main sources of agglomeration economies; knowledge spillovers, interfirm transaction relationships, and labor pooling. By using inter-industry relationships in each agglomeration effect (e.g., input-output linkage, knowledge transfers, and sharing types of workers), we build indices of each agglomeration effect in the cluster by industry-level, and we decompose each effect.

We find the following results. First, localization economies actually improve firm-level productivity in Vietnam. Firms in the clustered areas have higher productivities. However, the localization economies do not improve the state-owned firms. This may be due to the existence of local protectionism. Second, urbanization economies improve productivity only for foreign-owned firms. State-owned and private firms do not benefit from urbanization economies. These results imply that agglomeration economies may not be fully effective in Vietnam, especially in urbanization economies.

When examining the decomposition of the agglomeration effects, we find that agglomeration economies through transactions works only for private firms. On the other hand, agglomeration economies through knowledge spillovers and labor pooling work for foreign-owned firms.

The rest of this paper is organized as follows. The next section describes the data. Section 3 describes our empirical strategy. Section 4 provides our main results. Finally, Section 5 concludes the paper.

2. Data

For this study, we use Vietnamese enterprise database. We also use data from the fourth (2012) Establishment Census conducted by the General Statistics Office (GSO). The Census covers all economic entities including enterprises (state, non-state and foreign-invested sectors), and "non-farm individual business establishments". The census was conducted on 1 April (with a reference date of 31 December 2011) for enterprises, and 1 July (with a reference date of 1 July) for non-farm individual business establishments. The country has 341,600 enterprises as of 31 December 2011, and 4.63 million non-farm individual business establishments as of 1 July 2012. The data we compute for this study include location (district level), ownership type (state, non-state, foreign-invested), establishment code, number of laborers, and value added of the enterprises. We restrict the sample to manufacturing firms.

To capture inter-industry relationships in the three sources of agglomeration economies (i.e., transaction, knowledge spillovers, and labor pooling), we use the following data. To capture inter-industry transaction relationships, we use the input-output table for 2012 provided by the GSO. The input-output table contains an inter-industry transaction matrix with 164 X 164 sectors. By using the concordance provided by the GSO, we convert it to International Standards of Industrial Classification (ISIC) four-digit level. For the inter-industry knowledge spillovers measure, we use the result of a questionnaire survey included in the Establishment Census. This survey shows the flow of technology transfer from the suppliers to the enterprises and vice versa. Finally, for the labor pooling measure, we use the Vietnamese Household Living Standard Survey for 2012, which provides information regarding the number of workers of each job occupation in each industry.

3. Methodology

The purpose of this paper is to estimate the effects of agglomeration on productivity. This section explains how to construct measures of agglomeration and productivity, and how to estimate the effects of agglomeration on productivity.

3.1. Measure of agglomeration

We apply the methodology of Mori and Smith (2013) to identify the industrial clusters. This methodology identifies each cluster as a convex combination of contiguous regions within which the density of economic activities (e.g., establishments) is relatively uniform.

Their method detects the clusters as the convex contiguous combination of regions and cluster scheme as one or more disjoint clusters. Specifically, detected cluster scheme $\mathbb{C}^*$ maximizes the Bayesian information criterion (BIC) among each candidate cluster scheme $\mathbb{C}$ that is one or more disjoint clusters, C_j, j= 1, … , k,:

$$BIC_{\mathbb{C}} = L_{\mathbb{C}}(\hat{P}_{\mathbb{C}}|x) - \frac{k_{\mathbb{C}}}{2} \ln n$$

where $L_{\mathbb{C}}(\hat{P}_{\mathbb{C}}|x) = \sum_{j=0}^{k_{\mathbb{C}}} n_j(x)\, ln\hat{p}_{\mathbb{C}}(j) + \sum_{j=0}^{k_{\mathbb{C}}} \sum_{r\in C_j} n_r\, ln \frac{a_r}{a_{C_j}}$.

BIC increases with the larger log-likelihood of $\hat{P}_{\mathbb{C}}$, the location probability of cluster scheme $\mathbb{C}$ for each region, given an observed location pattern x, whereas BIC decreases with the penalty term composed by $k_{\mathbb{C}}$, which expresses the number of clusters in a cluster scheme $\mathbb{C}$, and n, which expresses the number of establishments in the total area. The location probability of cluster scheme $\mathbb{C}$ for cluster j= 1, … , k, $p_{\mathbb{C}}(j)$ can be rewritten as $\hat{p}_{\mathbb{C}}(j) = n_j(x)/n$ where $n_j(x)$ expresses the sum of the number of establishments in cluster j = 1, … , k and n represents the total number of establishments in the whole region. In the above log-likelihood functions, $n_j(x)$ and n_r are related with a sequence of independent location decisions by individual establishments. Because each region is included as part of a certain cluster C_j, j= 1, … , k or the residual set of all non-cluster regions, the right hand-side of the log-likelihood function of the location probabilities (i.e. a probability that a randomly sampled establishment locates in a region within a certain cluster) expresses the law of total probability: the log-likelihood function can be divided into two parts, namely, the first term, which gives the location probabilities that n_j establishments locate in a cluster C_j, j= 1, … , k, and the second term, which gives the location probability that n_r establishments locate in region r in cluster C_j, j= 1, … , k, given that individual establishments choose their location completely randomly within each cluster. The location probability of region r under the condition that an establishment locates in a cluster C_j, j= 1, … , k, is a_r/a_{C_j}, where a_r expresses the economic area[1] in region r and a_{C_j} represents the economic area in a cluster C_j, j= 1, … , k.

[1] The economic areas are obtained from the Global Land Cover 2000 by the Joint Research Center of the European Commission. Omitting any area that is not suitable for economic activity, we calculated the economic area of each region using the shape files.

The second step is divided into two parts. The first is to find the essential clusters, which are the most significant clusters in terms of their incremental contributions to BIC under the condition that the sum of incremental contributions to BIC exceeds a certain proportion (λ) of $BIC_{\mathbb{C}}$, The last part of the second step is to find the smallest convex-solid set in the total area containing the set of essential clusters. Intuitively, a convex-solid set means that the set is connected and has not only no dents in its perimeter but also no internal cavities. The regions in the smallest convex-solid set can then be regarded as an essential containment.

We use the essential clusters as a measure of industrial cluster. On particular situation on the industrial clusters in Vietnam, Gokan et al. (2016) conduct a detailed analysis.

3.2. Measure of productivity

To measure the firm-level productivity, we assume firm-level production function as Cobb-Douglas functional form as follows:

$$lny_e = \beta_1 lnk_e + \beta_2 lnl_e + \varepsilon_e,$$

where y_e is value added, k_e is capital, and l_e is number of employments of firm e. We estimate this equation by OLS, and we use the residual ε_e as for the log of total factor productivity (TFP) in firm e.

The parameters of the production function may differ across industries. To address this issue, we estimate this equation by each industry (two-digit level in International Standards of Industrial Classifications).

3.3. Estimating agglomeration economies

Based on the measures described in the above subsections, we estimate the agglomeration economies by the following estimation equation:

$$ln\mathrm{TFP}_e = \alpha + \beta_1 \mathrm{agglom}_r + Z_{er}\delta + \varepsilon_{er}$$

where agglom_r is the variable that represents the agglomeration economy in region r in which firm e locates. The coefficient β_1 represents the agglomeration economies.

The agglomeration economies can be classified into two types, localization and urbanization economies. Localization economies improve the firm-level productivities through the agglomeration of firms within the industry (e.g., Glaeser, Kallal, Shainkman, and Shleifer, 1992). On the other hand, urbanization economies improve the firm-level productivity through the diversity of industries (e.g., Jacobs, 1969). This study calculates a variable that represents the effects in both types of agglomeration economies. For the variable for localization economies, we use a dummy variable that is one if the firm e's locating district r is detected as a cluster of industry i to which the firm e belongs. This is an indicator variable that indicates whether the firm locates in the cluster of their own industry or not. On the other hand, a variable for urbanization economies should represent the diversity of industries. The cluster detecting methodology enables us to define the degree of urbanization by the number of layers of clusters in different industries in a region. For region r, we can count the number of industries that have a cluster in the region r. The number of industries that have a cluster in a region represents the variety of clustered industries in the region, and can be considered as the urbanization index for the region.

To control for the unobserved heterogeneity across regions and industries, we include prefectural fixed effects and industrial fixed effects. For the classification of industry, we use the two-digit ISIC code.

3.4. Decomposition of agglomeration economies

To decompose the agglomeration economies, we construct an index for each agglomeration effect by assuming that an establishment receives agglomeration effects in a cluster of an industry to which the establishment belongs.

First, we build measures of inter-industry relationships. For the inter-industry transaction relationships, we use input-output linkage information. Following the approach of Ellison, Glaeser, and Kerr (2010), we use the Vietnam Input-Output Table for 2012. The input-output table shows the inter-industry transactions, and we can therefore estimate the strength of the input-output linkages by computing the input and output coefficients of the transaction matrix. First, we compute the proportion of total input that sector A purchases from sector B and vice versa and take the maximum of these input coefficients as a measure of input linkages. We then calculate the proportion of total output that sector A provides to sector B and vice versa and take the maximum of the output coefficients as a measure of output linkages. Finally, we take the maximum of the above two measures to produce a measure of input-output linkages or transaction relationships. It is expected that firms in industries that are highly linked through transaction relationships will be more likely to locate close to each other.

Second, for the knowledge spillovers measure, we follow the approach by Emma, Carol, and Finn (2012). The survey conducted by the GSO contains the information on technology transfer from the suppliers to the firms. First, the technology transfer variable is constructed by calculating the proportion of the firms that received technology transfer from their suppliers and weight this according to the number of employees in each firm. Second, a matrix of inter-industry technology transfer is constructed by multiplying the weighted technology transfer variables by input coefficients. Third, we take the maximum technology transfer between the two sectors (say from sector A to B and vice versa) to construct a measure of technology transfer from the suppliers to the firms. Furthermore, using the information on technology transfer from the firms to the customers and output coefficients, we construct a measure of technology transfer from the firms to the customers. Finally, we take the maximum of the above two measures to construct a measure of technology transfer or knowledge spillovers.

Finally, for the labor pooling measure, we use the information on the number of workers in each occupation in each industry to calculate the correlation coefficient between a pair of industries and use this as a measure of labor pooling between industries. It is expected that firms in industries with similar skill sets will be more likely to locate jointly.

Using above inter-industry relationship measures, we construct the measure of agglomeration economies of type $t \in \{transaction,\ knowledge\ spillovers, labor\ pooling\}$ received by firm e in industry i locating region r as follows.

$$AGG_{e(i)r}^{t} = \sum_{j} d_{jr}\, w_{ij}^{t},$$

where d_{jr} is a cluster dummy that equals one if region r is detected as a cluster of industry j, and w_{ij}^{t} is the strength of inter-industry relationships on type t between industries i and j.

Using the above measures of sources of agglomeration economies, we estimate the following equation,

$$ln\mathrm{TFP}_{e(i)} = \alpha + \sum_{t} \beta^{t} AGG_{e(i)r}^{t} + Z_{er}\delta + \varepsilon_{er}.$$

We also control for industry and province fixed effects.

4. Results

4.1. Descriptive statistics and distributional analysis

First, we show the descriptive statistics in Table 1. Our data covers 42,389 observations.

Average value added is 19,459 million VND (approximately 871,982 USD).

Table 1

As we have mentioned, there are three types of firms in Vietnam, state-owned, private, and foreign firms, and the productivities of firms in each type may be different. We show the difference in the TFP distributions in these three types of firms in Figure 1.

Figure 1

As Figure 1 shows, the distributions of TFP are not so different across types of firms.

We then move to the investigation of the agglomeration economies. To investigate the localization economies we compare the productivity distributions between firms in clusters and non-clusters within an industry. Figure 2 shows the kernel density distributions of TFP in cluster and non-cluster areas.

Figure 2

The red line represents the kernel density distribution of the TFP of firms that locate in non-cluster areas, while the blue line represents that of firms in cluster areas. We can observe that the peak of the distribution has shifted rightward. This suggests the existence of localization economies. Furthermore, variation becomes wider on both sides in the distribution of clustered areas. This implies that the lowest productivity firms can survive only in the clusters. This is the opposite result to the literature that emphasizes the selection mechanism in the clusters (e.g., Arimoto, Nakajima, and Okazaki, 2014).

Next, we investigate urbanization economies. We define urbanized areas by the number of industries that have clusters in a region. Specifically, in each district, we calculate the number of industries that have clusters in the district, and we define urbanized areas if a district has the number of clustered industries above the median. Figure 3 shows the kernel density distributions of TFP in urbanized and non-urbanized areas. The red line represents the kernel density distribution of TFP of firms that locate in non-urbanized areas, and the blue line represents that in urbanized areas. The peak of the distribution shifts rightward and the variation becomes wider in higher productivity sides in the distribution of urbanized areas. This suggests the existence of urbanization economies.

Figure 3

The effects of agglomeration economies may differ between types of firms, and so we draw TFP distributions of urbanized and non-urbanized areas across types of firms. Figure 4 shows the kernel density distributions of state-owned firms' TFP in urbanized and non-urbanized areas. The red line represents the kernel density distribution of TFP of firms that locate in non-urbanized areas, and the blue line represents that in urbanized areas. The peak of the distribution is not different between the two distributions. Furthermore, the left tail of the distribution is longer in the firms in urbanized areas. This implies that low-productivity state-owned firms survive in urbanized areas.

Figure 4

Figure 5 shows the kernel density distributions of private firms' TFP in urbanized and non-urbanized areas. The red line represents the kernel density distribution of TFP of firms that locate in non-urbanized areas, while the blue line represents that in urbanized areas. The peak of the distribution is shifted rightward and the right tail of the distribution is longer in the distribution of urbanized areas. This implies that urbanization economies work for private firms.

Figure 5

Finally, Figure 6 shows the kernel density distributions of foreign owned firms' TFP in urbanized and non-urbanized areas. The red line represents the kernel density distribution of TFP of firms that locate in non-urbanized areas, and the blue line represents that in urbanized areas. The peak of the distribution is strongly shifted rightward in the distribution of urbanized areas. This implies that urbanization economies are particularly effective for foreign owned firms.

Figure 6

Table 2 shows descriptive statistics in TFP across types of firms and locations. There is no difference in TFP of state-owned firms between urban and non-urban areas. On the other hand, private and foreign firms' TFPs differ significantly between urban and non-urban areas. These results imply that agglomeration economies mainly work for private and foreign firms.

Table 2

Finally, the urbanization economies may differ in industries. Table 3 shows the descriptive statistics in TFP across industries. In 11 out of 22 industries in total, we find a statistically significant difference between urbanized and non-urbanized areas. The strength of the urbanization economies differs between industries. Most of the industries which have significant urbanization economies are light industry such as food, tobacco, textiles, and so on. This contrasts with the result of Henderson (2003), which finds urbanization economies in the high-tech sector. In Vietnam, the high-tech sector is still developing and urbanization economies may come from large demand in the urban areas.

Table 3

4.2. Results on agglomeration economies

In the descriptive statistics, we find both localization and urbanization economies, with the strength of these effects differing across types of firms. This subsection formally tests these effects using a regression analysis.

Table 4 shows the results of the estimation of the localization economy. Column (1) shows the results with all observations. The coefficient for the localization dummy is positive and significant. This implies that there is a localization economy that improves the productivity of firms located in clusters. We then conduct the same regression for each type of firms. Column (2) shows the results for state-owned firms. The coefficient for the localization dummy is positive but not significant. For state-owned firms, firm-level productivity is unchanged between firms in clusters and non-clusters. Localization economies do not work for the state-owned firms. A possible interpretation of the result is the local protectionism observed by Lu and Tao (2009) in China. They find that a high share of stated-owned firms' employment in a region impedes industrial agglomeration. Similarly, the existence of a state-owned firm impedes the agglomeration in the region reducing the effect of localization economies. Column (3) shows the results for private firms. The coefficient for the localization dummy is positive and significant. Localization economies function for private firms. Finally, Column (4) shows the results for foreign-owned firms. Similar to the results for private firms, the coefficient for the localization dummy is positive and significant, but the magnitude is much larger than that for private firms. Foreign firms therefore benefit more from localization economies.[2]

Table 4

We then move to an investigation of urbanization economies. The results are shown in Table 5. Column (1) shows the results for all the observations. The coefficient for the urbanization variable is positive, but not significant. This implies that urbanization economies do not improve firm-level productivity in Vietnam. This weak urbanization effects is consistent with the empirical research on urbanization economies (e.g., Henderson, 2003). Column (2) shows the result for the state-owned firms. The coefficient for the urbanization variable is negative but not significant. Column (3) shows the result for private firms. The coefficient for the urbanization variable here is positive but not significant. Finally, Column (4) shows the result for foreign-owned firms. The coefficient for the urbanization variable is positive and significant at the 1% level. In Vietnam, urbanization economies only improve the productivity of foreign-owned firms.

Table 5

4.3. Decomposition of agglomeration economies

Results of the analysis of the decomposition of agglomeration economies are shown in Table 6. Column (1) is the results with all observations. The coefficient for transaction is positive and significant. That is, firms in clusters of industries closely related through transaction relationships have higher productivity. The other agglomeration economies are not significant.

We then conduct the same regression for each type of firm. Column (2) shows the results for state-owned firms. Any agglomeration effects are not significant. This implies that state-owned firms do not benefit from any of the different types of agglomeration economies. This is consistent with the above baseline analysis that observes no localization and urbanization economies for state-owned firms. Column (3) shows the results for private firms. Similar to the baseline analysis (Column 1), the coefficient for transaction is positively significant. This implies that agglomeration

[2] We must also address the issue of causality. This result may come from the self-selected location choice of the firms. That is, foreign firms may be more sensitive than other firms to location, while highly productive firms would tend to locate clusters.

economies for private firms are mainly through inter-firm transactions. Finally, Column (4) shows the results for foreign-owned firms. This has totally different patterns to the other firms. Coefficients for labor pooling and knowledge spillovers are significantly positive. Agglomeration economies for foreign-owned firms come from knowledge spillovers and labor pooling. Foreign-owned firms use more advanced technology relative to the Vietnamese local firms, and it is natural that knowledge spillovers are beneficial for foreign-owned firms. In terms of labor pooling, we find that foreign-owned firms employ large numbers of workers in each plant. Thus the thick labor market of specialized workers would be crucial for their profitability.

5. Conclusion

This paper examines the agglomeration economies on firm-level productivity in Vietnam. By using Vietnamese firm-level data and the cluster detection method proposed by Mori and Smith (2013), we estimate the effects of agglomeration economies on firm-level productivity. Specifically, we consider the different effects of agglomeration economies between localization and urbanization and across types of firms; state-owned, private, and foreign-owned firms.

We find the following results. First, localization economies actually improve firm-level productivity in Vietnam. Firms in the clustered areas have higher productivities. However, localization economies do not improve state-owned firms. This may represent the existence of local protectionism. Second, urbanization economies improve productivity only for foreign-owned firms. State-owned and private firms do not benefit from urbanization economies. These results imply that the agglomeration economies may not be fully effective in Vietnam, especially in urbanization economies. The weak urbanization economies are consistent with literature such as Henderson (2003) that only finds urbanization effects in high-tech industries.

In transition economies, state-owned firms still have a large presence. In Vietnam, these firms do not necessarily have lower productivity than other firms. However, these firms do not benefit from agglomeration economies both in localization and urbanization. Additionally, the presence of foreign-owned firms in Vietnam has recently been increasing. These firms benefit greatly from both localization and urbanization economies. Furthermore, our results suggest that the sources of these agglomeration economies are knowledge spillovers and labor pooling, whereas private firms benefit from agglomeration economies through transactions.

References

[1] Arimoto, Y., K. Nakajima, and T. Okazaki (2014) "Sources of Productivity Improvement in Industrial Clusters: The Case of the Prewar Japanese Silk-reeling Industry," *Regional Science and Urban Economics* 46, pp. 27-41.

[2] Ciccone, A. (2002) "Agglomeration Effects in Europe," European Economic Review 46, pp. 213-227.

[3] Ciccone, A. and Hall, R. (1996) "Productivity and the Density of Economic Activity," American Economic Review 86, pp. 54-70.

[4] Drucker, J. and E. Feser (2012) "Regional Industrial Structure and Agglomeration Economies: An Analysis of Productivity in Three Manufacturing Industries," *Regional Science and Urban Economics* 42, pp. 1-14.

[5] Ercole, R. (2013) "Spatial Agglomerations of Two-Digit Manufacturing Industries within and across Vietnamese Locations," mimeo.

[6] Ellison, G., E. Glaeser, and W. Kerr (2010) "What Causes Industry Agglomeration?" *American Economic Review* 100(3), pp. 889-927.

[7] Glaeser, E., H. Kallal, J. Shainkman, and A. Shleifer (1992) "Growth in cities," *Journal of Political Economy* 100, pp. 1126-1152.

[8] Gokan, T., I Kuroiwa, K. Nakajima, and S. Sakata (2016) "Spatial Patterns of Manufacturing Clusters in Vietnam," mimeo.

[9] Henderson, V. (2003) "Marshall's Scale Economies," *Journal of Urban Economics* 53, pp. 1-28.

[10] Howard, E., C. Newman, and F. Tarp (2012) "Measuring Industry Agglomeration and Identifying the Driving Forces," WIDER Working Paper, No. 2012/84.

[11] Jacobs, J. (1969) *The Economy of Cities*, Vintage, New York.

[12] Lin, H., H. Li, and C. Yang (2011) "Agglomeration and Productivity: Firm-level Evidence from China's Textile Industry," *China Economic Review* 22, pp. 313-329.

[13] Lu, J. and Z. Tao (2009) "Trends and Determinants of China's Industrial Agglomeration," *Journal of Urban Economics* 65, pp. 167-180.

[14] Marshall, A. (1920) *Principles of Economics*, Macmillan, London.

[15] Mori, T. and T. Smith (2013), "A Probabilistic Modeling Approach to the Detection of Industrial Agglomerations," *Journal of Economic Geography* 14, pp.547–588.

[16] Rosenthal, S. and W. Strange (2001) "The Determinants of Agglomeration," *Journal of Urban Economics* 50(2), pp. 191-229.

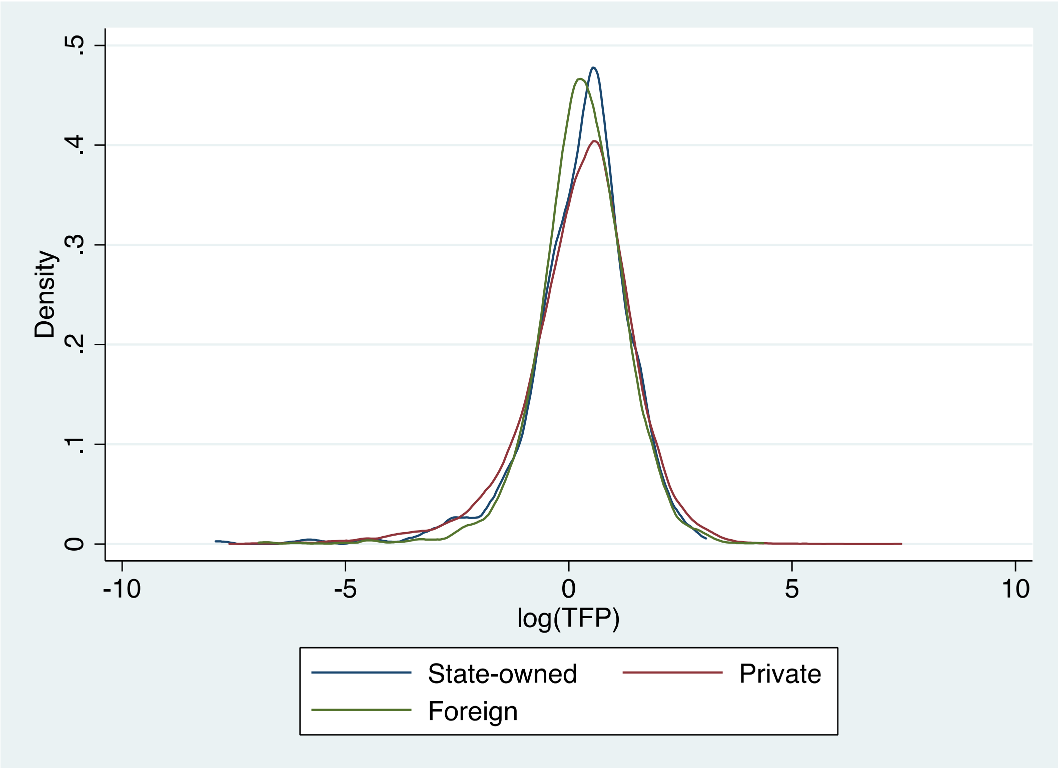

Figure 1 Kernel density distributions of TFP in three types of firms

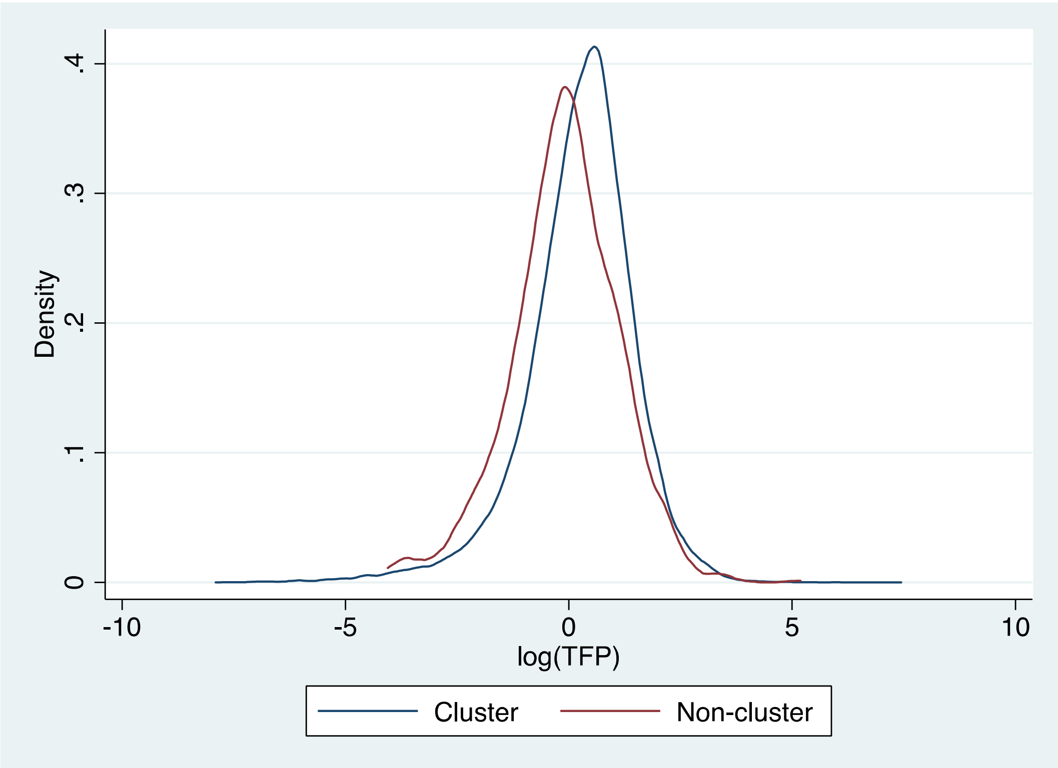

Figure 2 Kernel density distributions of TFP in cluster and non-cluster areas

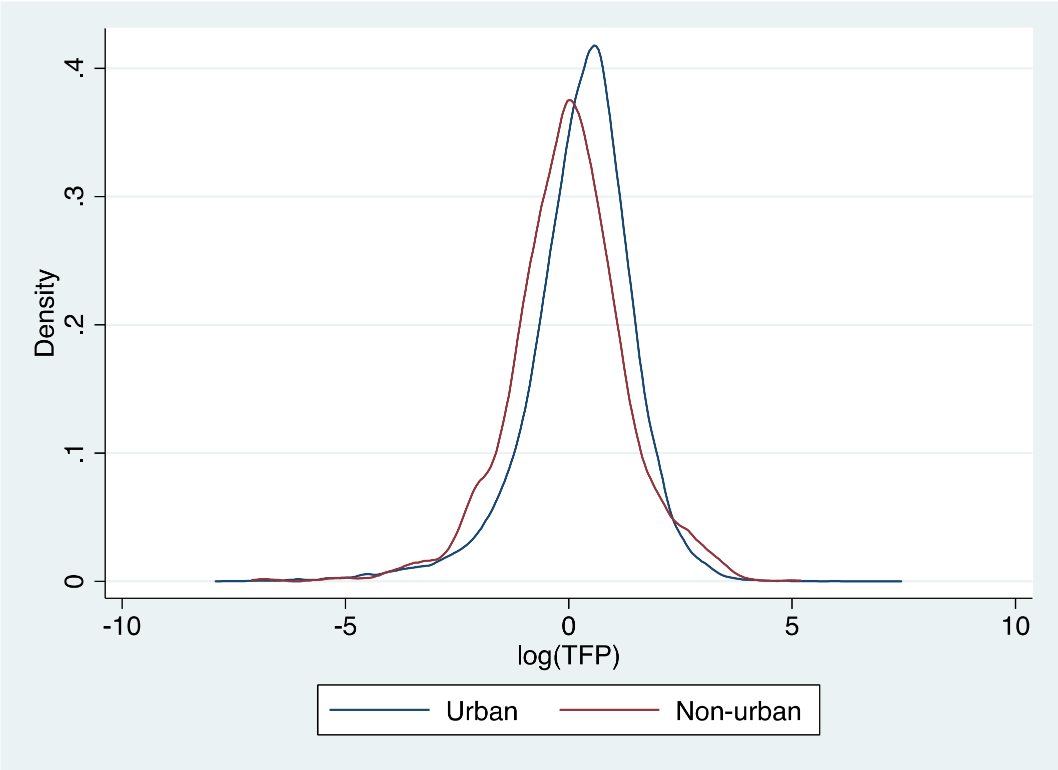

Figure 3 Kernel density distributions of TFP in urbanized and non-urbanized areas

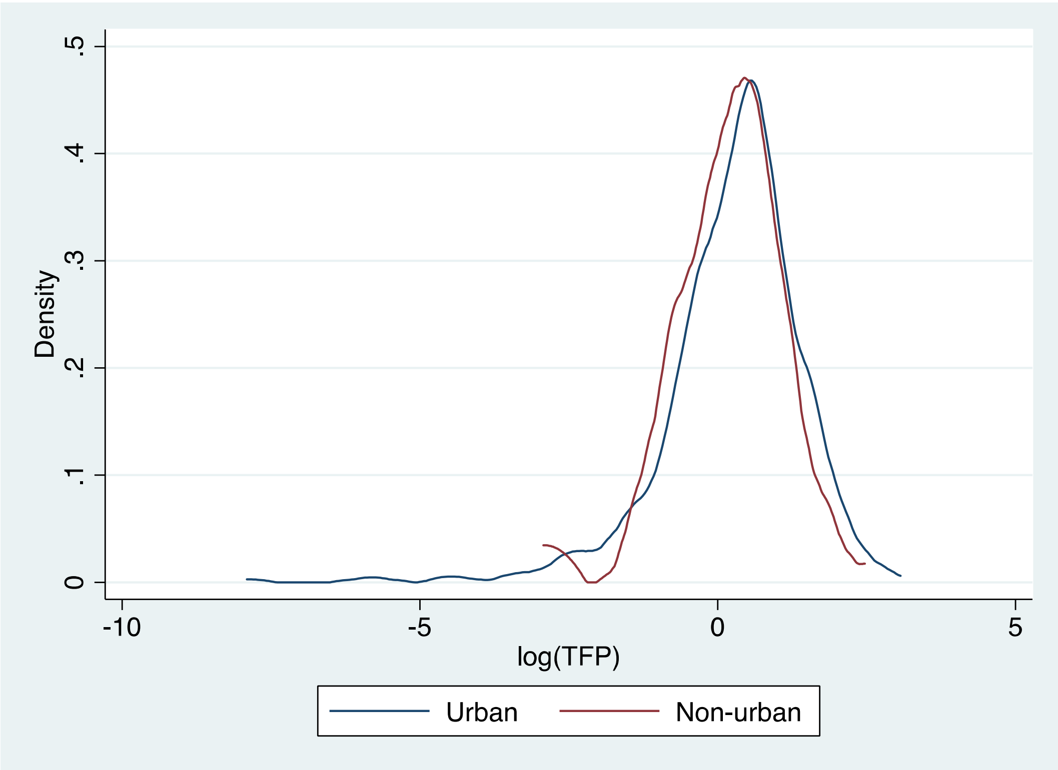

Figure 4 Kernel density distributions of TFP of state-owned firms in urban and non-urban areas

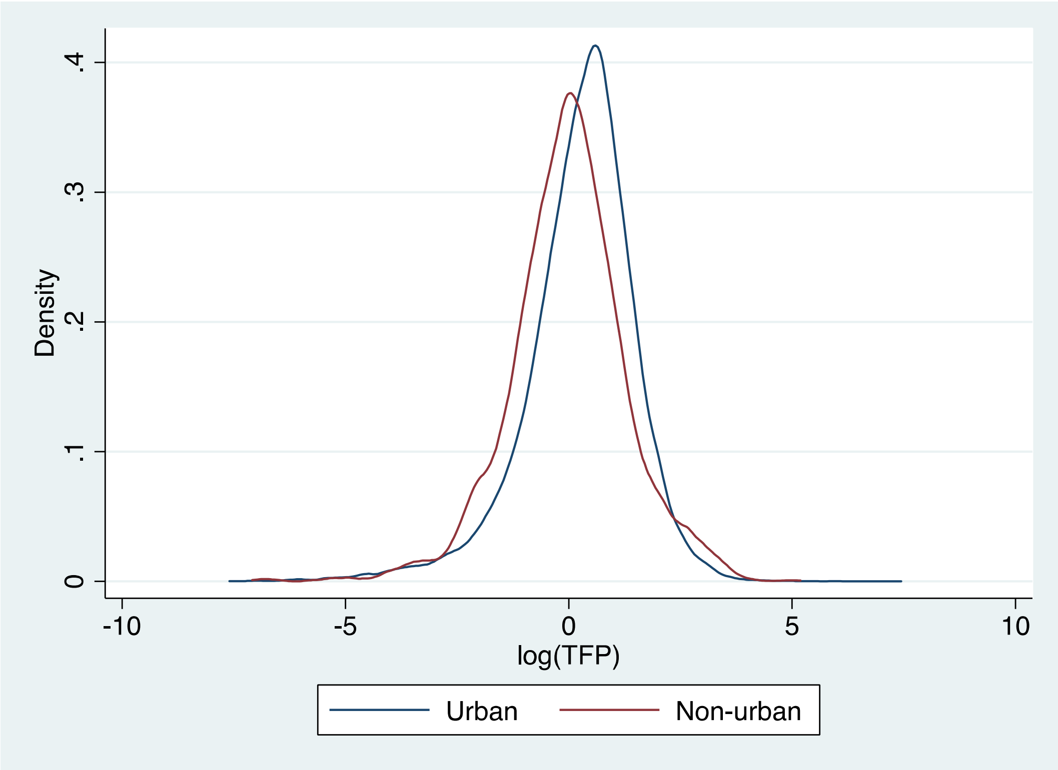

Figure 5 Kernel density distributions of TFP of private firms in urban and non-urban areas

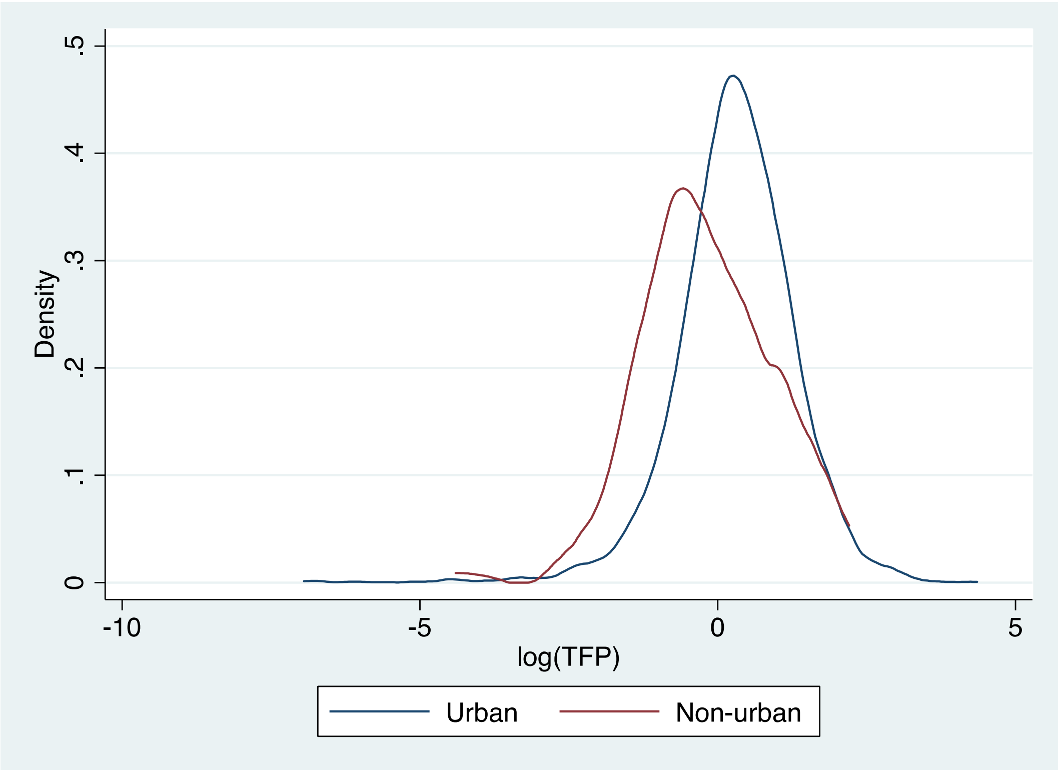

Figure 6 Kernel density distributions of TFP of foreign firms in urban and non-urban areas

Table 1 Descriptive statistics

Variables	No of obs.	Mean	SD
Value added	42289	19458.79	291359
Number of employments	42289	102.392	652.4412
Capital	42289	50764.5	470313.1
ln(TFP)	42289	0.237665	1.222124

Table 2 TFP in urban and non-urban areas: types of firms

	Urban			Non-urban				
Types	Obs.	Mean	SD	Obs.	Mean	SD	p-value	p<0.05
State-owned	549	0.255	1.181	63	0.161	0.958	0.552	
Private	41229	0.255	1.238	3840	−0.018	1.307	0.000	*
Foreign	4897	0.298	1.006	111	−0.174	1.119	0.000	*

Table 3 TFP in urban and non-urban areas

	Urban			Non-urban				
ISIC	Obs.	Mean	SD	Obs.	Mean	SD	p-value	p<0.05
10	4057	0.158	1.423	1247	0.385	1.377	0.000	*
11	1707	0.102	1.263	231	-0.01	1.303	0.252	
12	25	1.593	1.112	0				
13	2069	0.115	1.192	53	-0.551	1.258	0.000	*
14	4184	-0.109	1.161	144	-0.488	0.892	0.002	*
15	1168	-0.072	1.209	30	-0.314	0.848	0.361	
16	3168	0.302	1.315	882	-0.269	1.25	0.000	*
17	1797	0.552	1.026	42	-0.265	1.194	0.000	*
18	3438	0.662	1.105	38	0.301	0.9	0.074	
19	81	0.759	1.108	4	-0.675	2.236	0.021	*
20	1924	0.336	1.265	82	-0.277	1.541	0.000	*
21	336	0.564	1.054	7	0.039	0.851	0.192	
22	3265	0.462	1.191	42	0.143	0.985	0.126	
23	3079	0.205	1.147	546	-0.195	1.081	0.000	*
24	947	0.67	1.235	29	0.227	0.934	0.065	
25	7754	0.278	1.127	292	-0.034	1.177	0.000	*
26	609	0.015	1.187	6	-0.667	0.704	0.200	
27	1037	0.22	1.124	2	-0.739	1.737	0.229	
28	1173	0.469	1.057	22	0.028	1.243	0.073	
29	346	0.439	1.117	3	-0.808	2.126	0.118	
30	542	0.05	1.167	43	-0.311	0.97	0.075	
31	2853	0.141	1.156	235	-0.47	1.249	0.000	*
32	1116	0.016	1.245	34	-0.534	1.66	0.028	*

Table 4 Estimation of localization economy

	(1)	(2)	(3)	(4)
Localization dummy	0.155***	0.153	0.152***	0.367**
	(3.750)	(0.575)	(3.551)	(2.570)
Industry fixed effects	yes	yes	yes	yes
Province fixed effects	yes	yes	yes	yes
Sample	All	State-owned	Private	Foreign-owned
Observations	42,289	587	37,122	4,580
R-squared	0.056	0.323	0.057	0.130

Robust t-statistics in parentheses

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

Table 5 Estimation of urbanization economy

	(1)	(2)	(3)	(4)
Urbanization dummy	0.00123	-0.00114	0.00120	0.00274***
	(1.183)	(-0.424)	(1.014)	(3.126)
Industry fixed effects	yes	yes	yes	yes
Province fixed effects	yes	yes	yes	yes
Sample	All	State-owned	Private	Foreign-owned
Observations	42,289	587	37,122	4,580
R-squared	0.056	0.323	0.057	0.131

Robust t-statistics in parentheses

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

Table 6 Decompositions of agglomeration economies

	(1)	(2)	(3)	(4)
Transactions	0.0290***	0.0237	0.0347***	−0.00964
	(3.318)	(0.385)	(3.671)	(−0.852)
Knowledge spillovers	−0.0549	−0.108	−0.0767	0.145**
	(−0.984)	(−0.392)	(−1.252)	(2.339)
Labor pooling	0.00244	0.00123	0.00224	0.00442**
	(1.436)	(0.239)	(1.197)	(2.077)
Industry fixed effects	yes	yes	yes	yes
Province fixed effects	yes	yes	yes	yes
Sample	All	State-owned	Private	Foreign-owned
Observations	42,286	587	37,119	4,580
R-squared	0.057	0.323	0.058	0.132

Robust t-statistics in parentheses

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

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN