Industrial Agglomeration and Firm Size: Evidence from China

Dongya Li and Yi Lu*
University of Hong Kong

This version, January 2009

Abstract

This paper, by using annual surveys of manufacturing firms from 1998 to 2005 in China, first documents a positive correlation between industrial agglomeration and firm size, which is previously found in developed economies. Next, by using the system GMM and instrumental variable estimations, we identify that industrial agglomeration has a positive and statistically significant causal impact on firm size. Finally, we find that firms are more likely to benefit from locating with a number of large firms rather than with a large number of firms.

Keywords: Industrial Agglomeration, Firm Size, Agglomeration Economies, Urbanization Economies

JEL Codes: R12, L25, D21

*Corresponding author: Yi Lu, Faculty of Business and Economics, University of Hong Kong, Pokfulam Road, Hong Kong. Tel: (852) 2219 4220. Fax: (852) 2858 5614. Email: xzhongkanson@business.hku.hk. Lu thanks University of Hong Kong for the financial support.

1 Introduction

Industrial activities are unevenly distributed across space, e.g., *manufacturing belt* in the United States (Fritz, 1943), *blue banana belt* in the European Union (Delamaide, 1994), and *Pacific coast industrial belt* in Japan (Kitamura and Yada, 1977).[1] The agglomeration of industrial activities has significant impacts on firm behavior and firm performance such as productivity (e.g., Ciccone and Hall, 1996; Henderson, 2003), organization of production processes (e.g., Holmes, 1999; Li and Lu, 2009), and innovation (e.g., Feldman and Audretsch, 1999; Carlino, Chatterjee, and Hunt, 2007).

In two seminal papers, Kim (1995) and Holmes and Stevens (2002), by using plant-level data in the United States, find a positive correlation between industrial agglomeration and plant size both across and within manufacturing industries. Subsequent studies further confirm this finding by using datasets from other developed economies, e.g., Ireland (Barrios, Bertinelli, and Strobl, 2006) and Italy (Lafourcade and Mion, 2007).[2] An interesting research question is that does the same pattern exist also in developing economies, where economic environments differ a lot from their counterparts in developed economies. And more importantly, does the positive linkage between industrial agglomeration and firm size implies that firms become larger by locating in concentrated industrial areas or reflects the self-selection by larger firms into these areas? And how does industrial agglomeration affect firm size? Answers to these questions have important implications for both academic researches and government policies.[3] However, very few empirical studies have examined these issues. In this paper, we fill in the void by investigating empirically the impact of industrial agglomeration on firm size in China.

China presents a good setting to study this topic. Before 1978, China adopted a central planning system, and nearly all economic activities including location choice and production scale were determined by the central government, which was largely influenced by political considerations.[4]

[1] For more detailed description of spatial distribution of economic activities in the United States and Canada see Holmes and Stevens (2004a); in the European Union see Combes and Overman (2004); and in Japan and China see Fujita, Mori, Henderson, and Kanemoto (2004).

[2] Several theories have been proposed to explain the positive correlation between industrial agglomeration and plant size, e.g., Holmes and Stevens (2004b) and Wheeler (2008).

[3] For example, if industrial agglomeration has a causal impact on firm size, it not only contributes to the studies of agglomeration economies as proposed by Marshall (1890), Arrow (1962) and Romer (1986), but also suggests that government policies of setting up industrial zones to attract investments have their merits.

[4] Lu and Tao (2009) provides an example illustrating this point: in the late 1960s,

Confronted with the poor economic performance, however, China's central government started to reform its economy by gradually introducing private ownership and market competition in the late 1978. The economic reform not only induced a massive entry of privately-owned enterprises and foreign multinationals, but also restored incentives and decision-making powers in state-owned enterprises. As a result, there have been significant changes in the distribution of industrial activities. For example, Bai, Du, Tao, and Tong (2004) find a upward trend of industrial agglomeration in the latter half of 1985-1997, while Lu and Tao (2009) show that the industrial agglomeration continues during the period of 1998-2005. The fast-changing economic environment in China allows us to examine the interaction between industrial agglomeration and firm behavior.

Our dataset comes from annual surveys of manufacturing firms conducted by the *National Bureau of Statistics of China* for the period of 1998 to 2005. Using Holmes and Stevens (2002)'s specification, we find that firm size is positively and statistically significantly correlated with industrial agglomeration, which is consistent with the findings in the literature. Meanwhile, in terms of magnitude, the estimated coefficient in China (0.447) is very similar to that in the United States (0.436) as reported by Holmes and Stevens (2002).

To further investigate the relationship between industrial agglomeration and firm size, we follow Henderson (2003)'s estimation framework. Both the pooled OLS estimation and the panel estimation show that industrial agglomeration, measured by a firm's total neighboring employment within the same 4-digit industry and same region, has a positive and statistically significant impact on the firm's size. To identify whether industrial agglomeration has a causal impact on firm size, we use the system GMM estimation developed by Blundell and Bond (1998) and the instrumental variable estimation à la Li and Lu (2009). The system GMM estimation and the instrumental variable estimation results substantiate our early findings, showing that industrial agglomeration causes firms to become large in production scale. We next include two measures of urbanization economies as in Holmes (1999) in the regression analysis to ensure that our results are not driven by urbanization economies (Jacob, 1969). It is found that our findings are robust to inclusion of these two additional measures, though urbanization economies also cast a positive impact on firm size.

Finally, to investigate how industrial agglomeration affects firm size, we decompose our measure of industrial agglomeration into two parts: the num-

the production of key industrial products were re-allocated from coastal areas to interior regions as a preparation for possible wars with neighboring economies.

ber and the average size of a firm's neighboring firms. It is found that both the number and the average size of a firm's neighboring firms have positive and statistically significant impacts on the firm's size, whereas the latter has a much larger impact than the former. These results suggest that a firm is more likely to benefit from locating with a number of large firms than with a large number of firms.

In addition to the literature of agglomeration economies,[5] our paper is related to some recent studies about the determinants of firm size. For example, firm size is found to be positively related to financial intermediary development (Beck, Demirgüç-Kunt, and Maksimovic, 2006) and the quality of legal institutions (Kumar, Rajan, and Zingales, 2002; Beck, Demirgüç-Kunt, and Maksimovic, 2006; Laeven and Woodruff, 2007). While these papers study the impacts of various economic institutions, our focus here is the causal impact of industrial agglomeration on firm size.

The remainder of the paper is structured as follows. Section 2 describes data, and section 3 presents our empirical findings. The paper concludes with section 4.

2 Data

Our dataset comes from annual surveys of manufacturing firms conducted by the *National Bureau of Statistics of China* for the period of 1998 to 2005. These annual surveys cover all state-owned enterprises, and those non-state-owned enterprises with annual sales of 5 million RMB (Chinese currency) or more. And the dataset provides detailed information on firms' identification, operations and performance, including firm location, industry code and employment, which are of special interest to this study. The number of manufacturing firms with valid information (i.e., location code, 4-digit industry code and employment) varies from over 140,000 in the late 1990s to over 244,000 in 2005.

For our study, we need precise location and industry information of our sample firms. During the sample period, China's administrative boundaries and consequently its county, city or even region[6] codes have experienced some changes. Meanwhile, in 2003 a new classification system for industry codes was adopted in China to replace the old classification system that had been used from 1995 to 2002. To achieve consistency in the whole sample period

[5] See Rosenthal and Strange (2004, 2006) for reviews about the effects of agglomeration economies.

[6] Region here refers to 22 provinces, 4 province-level municipalities, and 5 minority autonomous regions in China.

(1998-2005), we convert the location codes and industry codes of all firms to those of year 1998. For more discussion on the details of adjustment, please see Lu and Tao (2009).

Moreover, firms in China may have branches located in regions other than its domicile, which may raise the concern of the multi-plants issue. However, according to the Article 14 of *The Company Law of the People's Republic of China*, "To set up a branch, the company shall file a registration application with the company registration authority, and shall obtain the business license." Thus, if a firm has branches that engage in business operations in regions other than its domicile, the *National Bureau of Statistics of China* collects each branch as a different observation in our dataset. For example, Beijing Huiyuan Beverage and Food Group Co., Ltd. has six branches, in Jizhong (Hebei Province), Youyu (Shanxi Province), Luzhong (Shandong Province), Qiqihar (Heilongjiang Province), Chengdu (Sichuan Province), and Yanbian (Jilin Province); the dataset accordingly counts them as six different observations belonging to six different regions, in addition to their parent company located in Beijing. As a result, we focus on region as our geographic unit.

3 Empirical Results

3.1 Benchmark

To give a first draw about the relation between industrial agglomeration and firm size, we follow Holmes and Stevens (2002)'s specification. Specifically, the measure for industrial agglomeration (the location quotient, $Q^x_{i,r}$) and the measure for firm size (the size quotient, $Q^x_{i,r}$) at the location-level are given by :

$$\begin{cases} Q^x_{i,r} = \frac{x_{i.r}/x_r}{x_i/x} \\ Q^s_{i,r} = \frac{x_{i.r}/n_{i.r}}{x_i/n_i} \end{cases} \quad (1)$$

where $x_{i.r}$, x_i, x_r and x are the total employment in industry i in region r, the total employment in industry i, the total employment in region r and the total employment respectively; and $n_{i.r}$ and n_i are the number of firms in industry i in region r and the number of firms in industry i respectively.

Columns 1-2 of Table 1 exhibit values of $Q^x_{i,r}$ and $Q^s_{i,r}$ for the whole example and each year, respectively. It shows that $Q^x_{i,r}$ experiences a modest decline in the late 1990s and then an steady increase in the early 2000s, which seems to contrast with the findings by Lu and Tao (2009) that the industrial agglomeration continues to increase for the same period. The difference could result from that the location quotient at the region level ($Q^x_{i,r}$) fails to take into account the large plant issue as pointed out by Ellison and Glaeser (1997). To partially address this issue, Holmes and Stevens (2002) propose another measures of the location quotient and the size quotient, which are at the firm-level. Specifically, the location quotient (Q_f^x) and the size quotient (Q_f^s) at the firm-level are given:

$$\begin{cases} Q_f^x = \frac{(x_{i.r}-x_f)/(x_r-x_f)}{(x_i-x_f)/(x-x_f)} \\ Q_f^s = \frac{x_f}{(x_i-x_f)/(n_i-1)} \end{cases} \tag{2}$$

where x_f is the employment of firm f in industry i in region r. As shown in Columns 4-5 of Table 1, Q_f^x steadily increases from 1.826 in 1998 to 2.008 in 2005, which is consistent with Lu and Tao (2009)'s findings.

The relation between industrial agglomeration and firm size is captured by the correlation between the location quotient (i.e., $Q_{i,r}^x$ and Q_f^x) and the size quotient (i.e., $Q_{i,r}^s$ and Q_f^s) (Holmes and Stevens, 2002):

$$\begin{cases} \beta^s = \frac{cov(q_{i,r}^s, q_{i,r}^x)}{var(q_{i,r}^x)} \\ \beta_f^s = \frac{cov(q_f^s, q_f^x)}{var(q_f^x)} \end{cases} \tag{3}$$

where lowercase q represents the natural logarithm of the uppercase counterpart. Column 3 and 6 of Table 1 show the estimated coefficients for β^s and β_f^s respectively. All the estimated coefficients are positive and statistically significant at the 1% level. Meanwhile, for the coefficients at the firm-level (β_f^s), the value continuously increases from 0.057 in 1998 to 0.093 in 2005.

These results imply that the positive correlation between industrial agglomeration and firm size previously found in developed economies also hold in developing economies such as China. Moreover, over the period of 1998-2005, the correlation between industrial agglomeration and firm size is found to become stronger and stronger in China. In the following subsections, we further investigate whether this relation is robust to the control of unobserved industry and region characteristics as well as urbanization economies, whether industrial agglomeration has a causal impact on firm size, and how industrial agglomeration affects firm size.

3.2 Unobserved Variables Issue

The above results regarding the positive correlation between industrial agglomeration and firm size could be biased due to some omitted variables such as industry and region characteristics.[7] To make sure the omitted variables do not bias our findings, we adopt Henderson (2003)'s estimation strategy. Specifically, we estimate the following equation:

$$size_{fir} = \alpha + \beta \cdot agglomeration_{fir} + X'_{fir}\gamma + \varepsilon_{fir} \tag{4}$$

where $size_{fir}$ is the logarithm of employment for firm f located in industry i in region r; $agglomeration_{fir}$ measures the degree of industrial agglomeration, which is the logarithm of firm f's total neighboring employment in the same 4-digit industry i and same region r; X'_{fir} is a set of control variables; and ε_{fir} is the error term. Standard error is clustered at the firm-level, to deal with the potential heteroskadasticity problem.

Two methods are used to estimate equation (4): the pooled OLS regression with a full set of industry, region, and year dummies,[8] and the panel estimation.[9] Regression results are reported in Table 2. As shown in Column 1, industrial agglomeration is positively and statistically significantly correlated with firm size. The estimated coefficient for *agglomeration* falls to 0.040 (shown in Column 2) when panel estimation is used but it is still positive and statistically significant.[10] These results are consistent with our previous findings that the industrial agglomeration has a positive impact on firm size, suggesting that our results are robust to the control for unobserved variables.

3.3 Endogeneity Issue

One may cast the doubt on the results in Tables 1-2 that whether it is industrial agglomeration causing the increase of firm size because it could be possible that large firms attract industrial agglomeration or choose to locate

[7] See Lu, Ni, and Tao (2008) for the discussion about the importance to deal with unobserved industry and region characteristics in estimating the agglomeration effect in China.

[8] We use 2-digit industry dummies rather than 4-digit industry dummies, because regressions with 1,406,355 observations and 517 4-digit industry dummies demand equipment that can do very intensive computing, to which we do not have access.

[9] Note that though the panel estimation is more efficient in controlling for time-invariant unobserved variables, it is more sensitive to measurement errors (e.g., Griliches and Hausman, 1986; Hauk and Wacziarg, 2006; Angrist and Pischke, 2009)

[10] The drop of estimated coefficient with panel estimation could result from the control for time-invariant firm unobserved characteristics that may correlate with both the degree of industrial agglomeration and firm size. Meanwhile, it also could be possible that much of variations regarding industrial agglomeration and firm size lies among the inter-firm rather intra-firm differences.

in agglomerated industrial areas.[11] To address these possible endogeneity issues and identify the causal impact of the industrial agglomeration on firm size, we use the system GMM estimation and instrumental variable estimation.

The system GMM estimation proposed by Blundell and Bond (1998) combines a set of first-differenced equations with level equations. It uses lagged first differences of endogenous variables as instruments in the level equation, while uses lagged levels of endogenous variables as instruments in the first-differenced equation. The validity of the system GMM estimation can be tested by the Sargan Test and the AR(2) Test (see Bond (2002), for more details). As shown in Column 1 of Table 3, the estimated coefficient of *agglomeration* remains positive and statistically significant. Meanwhile, both the Sargan Test and the AR(2) Test confirm that our system GMM estimation is valid.

We next use the instrumental variable estimation to address the endogeneity issues. Following Li and Lu (2009), we use the cross-region population of China in 1920 to instrument the degree of industrial agglomeration in contemporary China. The validity of our instrumental variable is based on two premises: the demand of a larger population attracts more manufacturers in each industry (see Krugman (1980) for a theoretical model, and Davis and Weinstein (2003) and Hanson (2005) for empirical supports), and the distribution of population persists over time (Davis and Weinstein, 2002). For more discussion on and various tests for the validity of this instrumental variable, please see Li and Lu (2009). The panel instrumental variable estimation results are reported in Column 2 of Table 3. The coefficient of the industrial agglomeration is positive and statistically significant; moreover, the value increases to 0.751, approximately 7 times larger than that from the pooled OLS estimation, suggesting that the estimation without controlling for the endogeneity could be biased downward.

Overall, results reported in Table 3 not only confirm but also substantiate our previous findings that the industrial agglomeration has a positive and statistically significant causal impact on firm size.

3.4 Urbanization Economies

Since the seminal work by Glaeser, Kallal, Scheinkman, and Shleifer (1992), there emerges a large literature regarding the relative importance of agglomeration economies and urbanization economies (e.g., Henderson, Kun-

[11] Some recent studies find that small firms contribute more to external economies (e.g., Henderson, 2003; Faberman, 2007; Glaeser and Kerr, 2008; Rosenthal and Strange, 2003, 2008).

coro, and Turner, 1995; Henderson, 1997, 2003; Quigley, 1998; Maurel and Sedillot, 1999; Combes, 2000; Gao, 2004). To rule out the concern that our findings could be driven by urbanization economies, we further include two measures related to urbanization economies as in Holmes (1999) in our analysis. The first one is the firm's total neighboring employment in the same region and same 2-digit industry but different 4-digit industry (denoted by $urbanization1$), and the second one is the firm's total neighboring employment in the same region but different 2-digit industry (denoted by $urbanization2$).

We stepwisely include these two urbanization measures and report regression results in Table 4. When only $urbanization1$ is included (Column 1), both the statistical significance and magnitude of the industrial agglomeration index do not change, though $urbanization1$ also has a positive effect on firm size. When $urbanization2$ is further included in the regression, the estimated coefficient for $agglomeration$ drops from 0.041 to 0.030 but it is still positive and statistically significant. These results suggest that although urbanization economies have positive impacts on firm size and explain some parts of the impact of industrial agglomeration on firm size, our main findings are not primarily driven by these urbanization economies.

3.5 Decomposition of Agglomeration Economies

As the last analysis, we study how industrial agglomeration affects firm size. To achieve this end, we decompose the degree of industrial agglomeration into two parts, the number and the average size of a firm's neighboring firms. Regression results are reported in Column 1 of Table 5. It is clear that both the number and the average size of a firm's neighboring firms have positive and statistically significant impacts on its size. However, the average size of the firm's neighboring firms have a much larger estimated coefficient (0.054) than that of the number of the firm's neighboring firms (0.034), suggesting that a firm is more likely to benefit from locating with a number of large firms rather than with a large number of firms.

In Column 2 of Table 5, we further include another four measures related to the decomposition of urbanization economies. Regression results exhibit a similar pattern that for both agglomeration economies and urbanization economies, firms are more likely to benefit from locating with a number of large firms rather than with a large number of firms.

4 Conclusion

Empirical studies using datasets from developed economies find that firm size is positively related to industrial agglomeration. In this paper, we attempt to investigate whether industrial agglomeration also has positive impacts on firm size in developing economies, where economic environments differ a lot from their counterparts in developed economies. Moreover, on top of the positive correlation, we aim at identifying whether industrial agglomeration has positive causal impacts on firms size and how industrial agglomeration affects firm size.

Using annual surveys of manufacturing firms from 1998 to 2005 in China, we first document a positive correlation between industrial agglomeration and firm size using both Holmes and Stevens (2002) and Henderson (2003)'s estimation specification. Second, by using the system GMM estimation and instrumental variable estimation, we find that industrial agglomeration has a positive and statistically significant causal impact on firm size. Third, through decomposition of industrial agglomeration, we find that firms are more likely to benefit from locating with a number of large firms rather than with a large number of firms.

References

Angrist, J.D. and J. Pischke. 2009. *Mostly Harmless Econometrics*, Princeton University Press.

Arrow, K.J. 1962. "The Economic Implications of Learning by Doing", *Review of Economic Studies* 29, 155-173.

Barrios, S., L. Bertinelli, and E. Strobl. 2006. "Geographic Concentration and Establishment Scale: An Extension Using Panel Data", *Journal of Regional Science* 46, 733-746.

Beck, T., A. Demirgüç-Kunt, and V. Maksimovic. 2006. "The Influence of Financial and Legal Institutions on Firm Size" Journal of Banking and Finance 30, 2995-3015.

Bond, S.R. 2002. "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice", *Portuguese Economic Journal* 1, 141-162.

Blundell, R. and S. Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics* 87, 115-143.

Carlino, G.A., S. Chatterjee, and R.M. Hunt. 2007. "Urban Density and the Rate of Invention", *Journal of Urban Economics* 61, 389-419.

Ciccone, A. and R.E. Hall. 1996. "Productivity and the Density of Economic Activity", *American Economic Review* 86, 54-70.

Combes, P.P. 2000. "Economic Structure and Local Growth: France, 1984-1993", *Journal of Urban Economics* 47, 329-355.

Combes, P.P. and H.G. Overman. 2004. "The Spatial Distribution of Economic Activities in the European Union", in J.V. Henderson and J.F. Thisse eds: *Handbook of Regional and Urban Economics*, Volumn 4, Amsterdan: Elsevier.

Delamaide, D. 1994. *The New Superregions of Europe*, New York: Dutton.

Ellison, G. and E.L. Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach", *Journal of Political Economy* 105, 889-927.

Faberman, R.J. 2007. "The Relationship between the Establishment Age Distribution and Urban Growth", working paper.

Feldman, M.P. and D.B. Audretsch. 1999. "Innovation in Cities: Science-based Diversity, Specialization and Localized Competition", *European Economic Review* 43, 409-429,

Fritz, W.G. 1943. "Natural Resource–Minerals", in: *Industrial Location and Natural Resources*, U.S. Government Printing Office, Washington, DC.

Fujita, M., T. Mori, J.V. Henderson, and Y. Kanemoto. 2004. "Spatial Distribution of Economic Activities in Japan and China", in J.V. Henderson and J.F. Thisse eds: *Handbook of Regional and Urban Economics*, Volumn 4, Amsterdan: Elsevier.

Gao, T. 2004. "Regional Industrial Growth: Evidence from Chinese Industries", *Regional Science and Urban Economics* 34, 101-124.

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

Glaeser, E.L., and W.R. Kerr. 2008. "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain", working paper.

Griliches, Z. and J.A. Hausman. 1986. "Errors in Variables in Panel Data", *Journal of Econometrics* 31, 93-118.

Hauk, W.R., and R. Wacziarg. 2006. "A Monte Carlo Study of Growth Regressions", *NBER* working paper #T0296.

Henderson, J.V. 1997. "Externalities and Industrial Development", *Journal of Urban Economics* 42, 449-470.

Henderson, J.V. 2003. "Marshall's Scale Economies", *Journal of Urban Economics* 53, 1-28.

Henderson, J.V., A. Kuncoro, and M. Turner. 1995. "Industrial-Development in Cities", *Journal of Political Economy* 103, 1067-1090.

Holmes, T. 1999. "Localization of Industry and Vertical Disintegration", *Review of Economics and Statistics* 81, 314-325.

Holmes, T. and J.J. Stevens. 2002. "Geographic Concentration and Establishment Scale", *Review of Economics and Statistics* 84, 682-690.

Holmes, T. and J.J. Stevens. 2004a. "Spatial Distribution of Economic Activities in North America", in J.V. Henderson and J.F. Thisse eds: *Handbook of Regional and Urban Economics*, Volumn 4, Amsterdan: Elsevier.

Holmes, T. and J.J. Stevens. 2004b. "Geographic Concentration and Establishment Scale: Analysis in an Alternative Economic Geography Model", *Journal of Economic Geography* 4, 227-250.

Kim, S. 1995. "Expansion of Markets and the Geographic Concentration of Economic Activities: the Trends in the U.S. Regional Manufacturing Structure, 1860-1987", *Quarterly Journal of Economics* 110, 881-908.

Kumar, K., R. Rajan, and L. Zingales. 2002. "What Determines Firm Size?", working paper.

Laeven, L. and C. Woodruff. 2007. "The Quality of the Legal System and Firm Size", *Review of Economics and Statistics* 89, 601-614.

Lafourcade, M. and G. Mion. 2007. "Concentration, Agglomeration, and the Size of Plants", *Regional Science and Urban Economics* 37, 46-68.

Li, B. and Y. Lu. 2009. "Geographic Concentration and Vertical Disintegration: Evidence from China", *Journal of Urban Economics*, forthcoming.

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

Lu, Y., J. Ni, and Z. Tao. 2008. "Growth in Cities, Revisited", working paper.

Kitamura, Y. and T. Yada. 1977. *Nippon-Kougyo No Chiiki-Kouzou* (in Japanese), Daimeido-Hakkou, Tokyo.

Marshall, A. 1890. *Principles of Economics*, London: Macmillan.

Maurel, F. and B. Sedillot. 1999. "A Measure of the Geographic Concentration in French Manufacturing Industries", *Regional Science and Urban Economics* 29, 575-604.

Quigley, J.M. 1998. "Urban Diversity and Economic Growth", *Journal of Economic Perspectives* 12, 127-138.

Romer, P.M. 1986. "Increasing Returns and Long-Run Growth", *Journal of Political Economy* 94, 1002-1037.

Rosenthal, S.S., and W.C. Strange. 2003. "Geography, Industrial Organization, and Agglomeration", *Review of Economics and Statistics* 85, 377-393.

Rosenthal, S.S., and W.C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies", in J.V. Henderson and J.F. Thisse eds: *Handbook of Regional and Urban Economics*, Volumn 4, Amsterdan: Elsevier.

Rosenthal, S.S., and W.C. Strange. 2006. "The Micro-empirics of Agglomeration Economies", in D. Arnott and P. McMillen eds: *A Companion* *to Urban Economics*, Wiley-Blackwell, Malden, MA.

Rosenthal, S.S., and W.C. Strange. 2008. "Small Establishments/Big Effects: Agglomeration, Industrial Organization, and Entrepreneurship", working paper.

Wheeler, C.H. 2008. "Worker Turnover, Industry Localization, and Producer Size", *Journal of Economic Behavior & Organization* 66, 322-334.

Table 1: Correlation between location quotient and size quotient

	1	2	3	4	5	6
	Location level			Firm level		
	$Q_{i,r}^{x}$	$Q_{i,r}^{s}$	β^{s}	Q_{f}^{x}	Q_{f}^{s}	β_{f}^{s}
All	1.573	1.055	0.447	1.949	1.014	0.081
1998	1.513	1.077	0.485	1.826	1.015	0.057
1999	1.490	1.069	0.482	1.829	1.016	0.073
2000	1.515	1.062	0.470	1.897	1.015	0.074
2001	1.531	1.060	0.456	1.949	1.017	0.075
2002	1.601	1.039	0.450	1.997	1.014	0.078
2003	1.641	1.029	0.436	1.982	1.013	0.082
2004	1.620	1.068	0.416	2.005	1.010	0.094
2005	1.676	1.033	0.399	2.008	1.011	0.093

Table 2: Pooled OLS and panel estimation results

	1	2
Estimation specification	Pooled OLS	Panel
Agglomeration	0.110***	0.040***
	(0.001)	(0.001)
Constant	3.105***	4.433***
	(0.020)	(0.011)
Year dummy	Yes	Yes
Industry dummy	Yes	-
Region dummy	Yes	-
Number of observation	1,406,355	1,406,355
R-squared	0.0956	0.0257
p-value for F-Test	0.0000	0.0000

Note: Standard errors, clustered at the firm-level, are reported in the parenthesis. *** indicates the statistical significance at the 1% level. In Column 2, overall R-squared is reported, whereas the within and between group R-squared are 0.0159 and 0.0293 respectively.

Table 3: System GMM and instrumental variable estimation results

Estimation specification	1 GMM	2 Panel+IV
Agglomeration	0.047***	0.751***
	(0.002)	(0.182)
Lagged size	0.634***	
	(0.005)	
Constant	1.376***	-2.095
	(0.032)	(1.678)
AR(2) Test	[12.596]***	-
Sargan Test	[2499.428]***	-
Year dummy	Yes	Yes
Industry dummy	-	-
Region dummy	-	-
Number of observation	931,661	1,404,341
R-squared	-	0.0268
p-value for chi2	0.0000	0.0000

Note: Standard errors, clustered at the firm-level, are reported in the parenthesis. *** indicates the statistical significance at the 1% level. The default hypothesis (H0) for AR(2) test is that there is no autocorrelation, while the default hypothesis (H0) for Sargan Test is that overidentifying restrictions are valid. In Column 2, overall R-squared is reported, whereas the between group R-squared are 0.0285.

Table 4: Urbanization economies

	1	2
Estimation specification	Panel	Panel
Agglomeration	0.041***	0.030***
	(0.001)	(0.001)
Urbanization 1	0.064***	0.036***
	(0.002)	(0.002)
Urbanization 2		0.352***
		(0.007)
Constant	3.683***	-1.137***
	(0.020)	(0.102)
Year dummy	Yes	Yes
Industry dummy	-	-
Region dummy	-	-
Number of observation	1,405,387	1,405,387
R-squared	0.0163	0.0028
p-value for F-Test	0.0000	0.0000

Note: Standard errors, clustered at the firm-level, are reported in the parenthesis. *** indicates the statistical significance at the 1% level. Overall R-squared is reported here, whereas the within and between group R-squared are 0.0181 and 0.0155 respectively for Column 1; and the within and between group R-squared are 0.0250 and 0.0022 respectively for Column 2.

Table 5: Decomposition of agglomeration economies

Estimation specification	1 Panel	2 Panel
Agglomeration—number	0.034***	0.023***
	(0.001)	(0.001)
Agglomeration—average size	0.054***	0.046***
	(0.002)	(0.002)
Urbanization 1—number		0.031***
		(0.002)
Urbanization 1—average size		0.058***
		(0.004)
Urbanization 2—number		0.346***
		(0.007)
Urbanization 2—average size		0.392***
		(0.011)
Constant	4.382***	-1.438***
	(0.013)	(0.106)
Year dummy	Yes	Yes
Industry dummy	-	-
Region dummy	-	-
Number of observation	1,406,355	1,405,387
R-squared	0.0405	0.0057
p-value for F-Test	0.0000	0.0000

Note: Standard errors, clustered at the firm-level, are reported in the parenthesis. *** indicates the statistical significance at the 1% level. Overall R-squared is reported here, whereas the within and between group R-squared are 0.0160 and 0.0467 respectively for Column 1; and the within and between group R-squared are 0.0255 and 0.0049 respectively for Column 2.