Trends and Determinants of China's Industrial Agglomeration

Lu, Jiangyong and Tao, Zhigang

Tsinghua University, University of Hong Kong

November 2007

Trends and Determinants of China's Industrial Agglomeration

Jiangyong Lu
Center for China in the World Economy,
and Department of Business Strategy and Policy,
Tsinghua University, Beijing, China, 100084
Email: Lujy3@sem.tsinghua.edu.cn

Zhigang Tao*
School of Business, University of Hong Kong,
Pokfulam Road, Hong Kong, China
852-2857 8223 (office), 852-2858 5614 (FAX), Email: ztao@hku.hk

November 2007

Abstract: This paper investigates trends and determinants of the spatial concentration of China's manufacturing industries using a large firm-level data for the time period of 1998 to 2005. It is found that the overall industrial agglomeration in China has increased steadily in recent years though it is still much lower than those of the well-developed market economies (such as United States, United Kingdom, and France). It is also found that local protectionism among China's various regions obstructs China's industrial agglomeration while Marshallian externalities facilitate the process of spatial concentration of manufacturing industries. On an optimistic note, there is evidence that the negative impacts of local protectionism have become less significant over time but those of Marshallian externalities are gaining in importance, which is consistent with the overall trend of China's industrial agglomeration.

Keywords: local protectionism, Marshallian externalities, industrial agglomeration
JEL classification codes: L11, R12, R30.

* We thank Editor Jan Brueckner and two anonymous referees for insightful comments and suggestions, and the University of Hong Kong and CCWE of Tsinghua University for financial support.

Trends and Determinants of China's Industrial Agglomeration

Abstract: This paper investigates trends and determinants of the spatial concentration of China's manufacturing industries using a large firm-level data for the time period of 1998 to 2005. It is found that the overall industrial agglomeration in China has increased steadily in recent years though it is still much lower than those of the well-developed market economies (such as United States, United Kingdom, and France). It is also found that local protectionism among China's various regions obstructs China's industrial agglomeration while Marshallian externalities facilitate the process of spatial concentration of manufacturing industries. On an optimistic note, there is evidence that the negative impacts of local protectionism have become less significant over time but those of Marshallian externalities are gaining in importance, which is consistent with the overall trend of China's industrial agglomeration.

Keywords: local protectionism, Marshallian externalities, industrial agglomeration
JEL classification codes: L11, R12, R30.

1. Introduction

Since 1978, China has undergone dramatic transformations from a centrally planned economy to a market economy. Along with this process, there have been significant changes in the geography of China's industrial activities. Before 1978, almost every economic activity including location choice was centrally planned, and those plans were not necessarily drawn according to market forces but rather influenced by political considerations. For example, in the late 1960s, there was a drive to relocate production of key industrial products from coastal areas to interior provinces in preparation for wars with neighboring countries and regions. With economic reform, it is expected that economic forces for industrial agglomeration should have redressed some of poor location choices of economic activities caused by central planning and played an important role in determining China's economic geography.

However, both anecdotal evidence and statistical analysis suggest that the same economic reform in China has led to the rise of local protectionism among China's various regions, which in turn slows down the process of industrial agglomeration as would have been determined by the economic forces. Indeed, a key policy introduced in China's economic reform is that of fiscal decentralization, under which the local governments can share tax revenue with China's central government. This policy is a double-edged sword, providing the local governments with strong incentives both for developing the local economy and for protecting local firms and industries from regional competition. Based on aggregated sectoral data and inter-regional input-output tables, Young [44] and Poncet [36] argue that local protectionism in China grew more and more serious over the 1990s. Meanwhile, Fan and Wei [16] find that both the pattern and the speed of price convergence in China are highly comparable with those measurements in well-developed market economies, and thus provide support for the view that the Chinese economy has been gradually transformed into a market-oriented economy. Using industry-level data, Bai, Du, Tao, and Tong [6] show the degree of industrial agglomeration in China first went down and then climbed up during the period of 1985-1997, and they also find that local protectionism has greater effects on industrial agglomeration than scale and external economies do during that sample period. Contrasting China's coastal area with its interior for the period of 1985-1994, Fujita and Hu [18] find that China's industrial production showed strong agglomeration toward the coastal area, and that income disparity between the two areas had been increasing.

The economic forces for industrial agglomeration, against those of local governments' incentives for protecting local firms and industries, make the study of China's economic geography exciting and challenging. In this paper, we use a large data set of China's manufacturing firms for the period of 1998 to 2005 to study the amalgam of economic and political forces shaping China's industrial agglomeration.

A critical theoretical condition for industrial agglomeration is the free flow of goods and services across regions without any political interference, but reality is often just the opposite. Local protectionism slows down the process of industrial agglomeration within a country, similar to the adverse impacts of national protectionist policies on international trade and specialization.[1] Despite their importance, studies on impacts of protectionist policies have been quite limited mainly due to the measurement difficulties. In this paper, we discuss the central and local government relations in China, and develop two indirect measures for local protectionism – local taxes to sales ratio and share of state-owned output, which are proxies for the incentive of local government officials to protect local firm and industries and allow us to examine the impacts of local protectionism on China's industrial agglomeration.

While our focus is on local protectionism, we also take into account the main economic forces favoring industrial agglomeration. First, it is the spillover effects, the study of which can be traced all the way back to Adam Smith [43]. There are three types of the spillover effects (Marshall [31], and henceforth *Marshallian externalities*), namely, knowledge spillovers, labor market pooling, and input sharing. Second, industrial agglomeration could be caused by regional variations in resource endowments, just like that of national variations in resource endowments on international specialization (Ohlin [34]). This force of industrial agglomeration, however, hinges upon the assumption of immobility of natural, physical, and human resources. Third, industrial agglomeration could also arise for those industries that exhibit increasing returns to scale (Krugman [28]).[2]

---

1 In recent years, research focus has been shifted towards political factors that may facilitate or obstruct the process of geographic concentration of economic activities. For example, Ades and Glaeser [1] show that political instability is associated with urban concentration. Holmes [22] classifies states in the United States as either pro-business or anti-business, and finds that the manufacturing share of total employment increases by about one-third when one crosses the border from an anti-business state into a pro-business state, which suggests that state policies matter in attracting businesses.

2 A fast-growing empirical literature has emerged and provided evidence on the validity of these micro-foundations of industrial agglomeration. Kim [27] and Ellison and Glaeser [14] examine the explanatory power of the resource endowment theory, Audretsh and Feldman [5] look into the importance of knowledge spillovers, Holmes [23] studies the role of input sharing, and Rosenthal and Strange [38] provide a comprehensive test of multiple determinants of agglomeration. See Rosenthal and Strange [40], and Duranton and Puga [13] for excellent surveys of recent empirical and theoretical studies on agglomeration economies.

The plan of the paper is as follows. In section 2, we first describe our data set from the Annual Survey of Industrial Firms conducted by China's National Bureau of Statistics for the period of 1998 to 2005, and then construct the Ellison and Glaeser [14] index (henceforth EG index) of China's industrial agglomeration following some of the latest developments in the economic geography literature. We find a consistently increasing time trend of industrial agglomeration in China for this period in contrast to some of the studies in the literature (Young [44]). However, comparisons with the EG indices of manufacturing industries in selected developed countries reveal that the degree of agglomeration in China's manufacturing industries remains considerably low despite its increasing trend.

In Section 3, we investigate the determinants of China's industrial agglomeration. Our focus is on local protectionism and Marshallian externalities, with controls for resource endowments and scale economies. Given the short time period of our data set, we first pool all the firms and years and run simple ordinary least square regressions on the correlation of the EG index and the variables for local protectionism, Marshallian externalities, resource endowments, and scale economies. To deal with the potential endogeneity problem, we use data from China's third industrial census conducted in 1995 to construct instrumental variables and re-run the regressions. The main findings from both the ordinary least square estimations and the instrumental variable approach are: (1) the two proxies for local protectionism – *local taxes to sales ratio* and *share of state-owned output* – have negative and statistically significant impacts on industrial agglomeration; (2) the variable for knowledge spillovers (*new products to output ratio*) and the variable for both knowledge spillover and input sharing (*exported output to total output ratio*) have positive and statistically significant impacts on industrial agglomeration; and (3) there is evidence for both scale economies and resource endowment.

As China has undergone dramatic changes during its economic reform, the importance of some of the determinants of industrial agglomeration may change over time. We thus split our sample into two sub-periods: 1998-2001 and 2002-2005, and re-run both the OLS and instrumental variable regressions. Indeed, it is found that the negative impacts of local protectionism have diminished in importance over time whereas the positive impacts of knowledge sharing have gained importance across the sample period. The paper concludes in Section 4 with some discussion for future work.

2. Trends of China's industrial agglomeration

2.1. Data

The main data set for this study comes from the Annual Survey of Industrial Firms (ASIF) conducted by China's National Bureau of Statistics for the period of 1998 to 2005. The survey covered all state-owned enterprises and those non-state-owned enterprises with annual sales of five million Renminbi (equivalent to US$ 660,000) or more in the following three types of industries: (1) mining, (2) manufacturing, and (3) production and distribution of electricity, gas and water. Table 1a shows the number of firms covered in the survey throughout the sample period: it ranges from 161,000 to 270,000. The location choice of firms in the first and third categories is heavily influenced by regional disparities in resource endowment. We thus focus on the sub-sample of manufacturing firms with the goal of investigating various determinants of industrial agglomeration. As shown in Table 1a, the number of manufacturing firms covered in the sample ranges from 146,000 to 251,000. There is a clear upward time trend, mainly because manufacturing firms in China have been growing rapidly over the sample period with more and more firms having annual sales of five million Renminbi or more. Following the literature (Ellison and Glaeser [14]; Rosenthal and Strange [38]), employment figures will be used to measure geographic concentration of manufacturing activities. As a result, those observations with missing or zero employment figures are deleted with a loss of less than 5% of the data (see Table 1a for details).

Table 1a here

There are two drawbacks with the Annual Survey of Industrial Firms dataset. First, it does not cover small non-state-owned enterprises. The estimation of industrial agglomeration could be biased, if the distribution of small non-state-owned enterprises varies across regions.[3] Second, the Annual Survey of Industrial Firms data are firm-level data, not plant-level data typically used in the literature. Multi-plant firms have become more common since China initiated its economic reform in 1978, and it may affect the analysis of industrial agglomeration if this phenomenon varies

[3] It is generally believed that, in choosing production locations, non-state-owned enterprises are less likely to be influenced by local government incentives and are more likely to be led by economic forces. The absence of small non-state-owned firms in the ASIF data implies that our analysis may under-estimate China's industrial agglomeration.

from regions to regions.[4]

There is an alternative dataset that is in principal the most ideal data set for studying industrial agglomeration: the industrial census of 1995. However, the 1995 industrial census data, unfortunately also a firm-level data, is of poor quality. The number of enterprises shrinks from 750,000 to 119,790 after the deletion of firms with missing value in total sales, number of employees, or fixed capital (Pan and Zhang [35]).[5] Thus in this study we use the Annual Survey of Industrial Firms as the main data source, and the 1995 industrial census data to construct instrumental variables for regression analysis. We will discuss possible future work using plant-level datasets and testing the robustness of our results obtained with the firm-level datasets.

Information on firm location is essential for studies of economic geography. For each firm in the ASIF data set, there is information on its address and the name of county, city and province where it is located. The existing studies have shown that the choice of geographic scope may affect the measure of industrial agglomeration – the so-called border effects. More recently, there have been several studies using ZIP code as the basic geographic unit or even using precise location data to minimize the border effects (Rosenthal and Strange [39]; Duranton and Overman [12]). However, the focus of this study is to examine the impacts of local protectionist policies on industrial agglomeration. Hence, the ideal geographic scopes for this study are the ones corresponding to the administrative areas where local government officials can have policies influencing inter-regional trade and industrial agglomeration. For this reason, county will be treated as the most disaggregated geographic scope, followed by city and province.

Along with China's spectacular economic growth, its administrative boundaries and consequently codes of counties, cities or even provinces have experienced significant changes in the last thirty years. For example, new counties could be established, while existing counties could be combined into larger ones or even elevated to cities. From 1998 to 2005, the number of counties in China increased from 2,496 to 2,862 (a total of 366), while the number of changes in county codes was 648. As firms may not be aware of the changes in the county codes, they may misreport in the annual surveys of industrial firms. Furthermore, even if the county codes reported are accurate, they may not be comparable across years. To address these problems, we first check the

---

[4] For manufacturing firms, firm location is generally the location for production, though there could be multiple locations for production.

[5] Since 1998, however, a direct reporting system has been adopted by China's National Bureau of Statistics, which has ensured the quality of statistical data (Holz [24]; Holz and Lin [25]). The 1998-2005 ASIF data is thus of much better quality than the 1995 census data.

accuracy of the county codes based on firms' reported addresses. Next, using the 1999 National Standard (promulgated at the end of 1998 and named GB/T 2260-1999) as the benchmark classification system of county codes, we convert the county codes of all firms to that benchmark system.

Aside from firm location, we also need information on firms' primary industry codes in order to conduct a study of China's industry agglomeration. For each firm in the ASIF data set, there is information on its primary 2-digit, 3-digit, and 4-digit industry codes. However, in 2003, a new classification system for industry codes (named GB/T 4754-2002) was adopted to replace the old classification system (named GB/T 4754-1994) that had been used from 1995 to 2002. To make the industry codes in the whole sample period (1998-2005) consistent, we convert the industry codes in the 2003-2005 data to the old classification system by using a concordance table (in the case of a new 4-digit code corresponding to an old 4-digit code, or several new 4-digit codes corresponding to an old 4-digit code) or by assigning a new code with an old code based on product information (in the case of several old 4-digit codes corresponding to a new 4-digit code). Industrial agglomeration will then be measured at the 4-digit industry level, followed by 3-digit and 2-digit industry levels with increasing industrial scope.

Tables 1b and 1c here

Tables 1b and 1c give the number of firms by industry and year and the size of employment by industry and year, respectively. The first column is the list of two-digit manufacturing industries covered in the survey, which are comparable to those of the Standard Industrial Classification codes. There is a general trend of relative decline in the first half of the sample period (1998-2001), especially in terms of the number of firms, presumably due to the negative impacts of the 1998 Asian Financial Crisis, and then a trend of robust growth in the remaining sample period. All but three industries have seen an increase in the number of firms, but about half of the industries have witnessed an increase in the size of employment. These results suggest increasing competition (more firms but with less employment) in China's manufacturing industries during the sample period.

2.2. Measuring China's industrial agglomeration

Most existing studies of China's industrial agglomeration have relied on highly aggregated data in terms of both industrial and geographic scopes. However, measures based on the aggregated data may fail to give an accurate picture of China's industrial agglomeration. For example, the car manufacturing industry and the bicycle manufacturing industry are two 4-digit industries of the same 2-digit transport industry, but they have different characteristics and exhibit different spatial patterns. Industrial agglomeration measured at the 2-digit industry level could then be misleading. Similarly, industrial agglomeration could be high when measured at the aggregated regional levels, but it is in fact low at the disaggregated regional levels, causing the so-called "Modifiable Area Unit Problem" (Arbia [3]). We deal with these problems by measuring industrial agglomeration at various industrial scopes (from 4-digit level to 3-digit level and then to 2-digit level) and geographic scopes (county, city and province) (Rosenthal and Strange [39]; Devereux, Griffith and Simpson [9]).

All existing studies of China's industrial agglomeration use the measurement of either the Gini or the Hoover index. However, as Ellison and Glaeser [14] point out, those coefficients do not take into account the impacts of industrial structure and may fail to give an accurate measure of industrial agglomeration. To address the problem, Ellison and Glaeser [14] construct a model-based index of geographic concentration (called γ index or EG index) which takes a value of zero if employment (or output) is only as concentrated as it would be had the plants in the industry chosen locations randomly by throwing darts at a map. The γ index takes the following form:

$$\gamma_i \equiv \frac{G_i - \left(1 - \sum_r x_r^2\right) H_i}{\left(1 - \sum_r x_r^2\right)\left(1 - H_i\right)}.$$

$G_i \equiv \sum_r (x_r - s_r)^2$ is the spatial Gini coefficient, where x_r is the share of total output of all industries in region r and s_r is the share of output for region r in industry i. $H_i \equiv \sum_i z_i^2$ is the Herfindahl index of industry i, with z_i standing for the output share of a particular firm in industry i. The Gini coefficient is expected to be larger in industries consisted of fewer and larger firms, even if locations were chosen completely at random (Dumais, Ellison and Glaeser [11]). The γ index is essentially the difference between G_i and H_i, measuring the degree of industrial agglomeration that is beyond the level implied by the industrial structures.

The compilation of γ requires the use of firm-level data of employment or output. With the firm-level data set from the Annual Survey of Industrial Firms, this paper represents the first attempt to measure China's industrial agglomeration by γ.[6] In principle, both employment data and output data can be used to calculate γ. Employment data is preferred to output data in the existing studies, as measurement using the output data may compound the impact of employment with that of capital.[7] In the case of China, however, there are concerns that employment data may suffer from the problem of surplus labor, which is especially severe in state-owned enterprises (see Bai, Du, Tao and Tong [6] for more discussion on the surplus labor problem). In this paper, for ease of comparison with the literature results, we use employment data to construct the EG index of China's industrial agglomeration.

2.3. Agglomeration of China's Manufacturing Industries

The γ indices are calculated at various geographic scopes (province, city and county) and industrial scopes (2-, 3-, and 4-digit industries). Weighted means (by employment) of γ_i indices across industries for each year of the sample period are given in Table 2a. Several interesting patterns can be found. First, similar to the findings of Rosenthal and Strange [38], the average level of agglomeration increases as one goes from 2 to 3-digit industries and from 3 to 4-digit industries, and it is also true as the geographic scope goes from county to city and from city to province.[8] Second, the γ_i indices for all possible combinations of industrial and geographic scopes have increased during the sample period of 1998-2005, which suggest increasing geographic concentration in China's manufacturing industries. These results are in contrast to the findings obtained using aggregate data by Young [44] that industrial agglomeration has decreased throughout China's economic reform.

Table 2a here
Table 2b here

---

[6] Using the EG index, Alecke, Alsleben, Scharr, and Untiedt [2], Devereux, Griffith and Simpson [9], Maurel and Sedillot [32], and Rosenthal and Strange [38] study spatial agglomeration of manufacturing industries in Germany, the U.K., France, and the U.S., respectively.

[7] A firm's output reported by China's National Bureau of Statistics measures its value added, and it is thus not influenced by how vertically integrated the firm is.

[8] See Rosenthal and Strange [38] for discussion of possible reasons behind the patterns.

Table 2b reports the EG index (γ_i), the corresponding Gini index (G_i), and Herfindahl index (H_i) calculated at the county level for all the 2-digit industries throughout the sample period.[9] Between 1998 and 2005, all but one industry (Garments & Other Fiber Products) had increasing γ indices.[10] Electronic & telecommunications had the biggest absolute increase in the γ index, followed by furniture manufacturing, and leather, furs, down & related products. Based on the γ indices of 2005, the three most aggregated industries are stationery, educational & sports goods; electronic & telecommunications; and leather, furs, down & related products, while the three least aggregated industries are metal products; papermaking & paper products; and printing & record pressing.

The EG index (γ_i) is in essence the difference between the Gini index (G_i) and the Herfindahl index (H_i), measuring the extent of geographic concentration that is beyond the level implied by the industrial structures. It is thus possible that industries with high Gini indices may have low EG indices, whereas industries with low Gini indices may have high EG indices. As shown in Table 2b, in 2005, tobacco processing ranked 20th among all industries in the EG index despite the fact its Gini index was the 3rd highest. It turns out that much of the industry's high Gini coefficient was due to its highly concentrated industrial structure (the highest Herfindahl index among all industries). A counter-example is nonmetal mineral products, which ranked 16th in the EG index despite the fact its Gini coefficient ranked 22nd of all industries in 2005. This is because much of the industry's low Gini coefficient was caused by its fragmented industrial structure (the lowest Herfindahl index among all 2-digit industries).

The γ index is designed to facilitate comparison across industries, across countries, and over time. The γ_i indices for manufacturing industries in the United Kingdom, the United States, and France have been studied by Devereux, Griffith and Simpson [9], Ellison and Glaeser [14], and Maurel and Sedillot [32], respectively, and their main findings are summarized in Table 2c together with ours. Note that these studies are carried out using data of various industrial and geographic scopes, and therefore the results are not directly comparable. The study of the U.S. manufacturing industries by Ellison and Glaeser [14] is the most comparable one to ours, given that similar

9 We choose the combination (2-digit industry level and county level) that gives the lowest EG index.

10 Among 171 3-digit industries, 145 industries have increasing γ_i from 1998 to 2005; among 540 4-digit industries, 404 industries see γ_i increased from 1998 to 2005.

geographic and industrial scopes are used.[11] Following the definitions of *not very concentrated industries*, *somewhat concentrated industries*, and *very concentrated industries* in Ellison and Glaeser [14], we find that 75.98%, 16.2%, and 7.82% of all 4-digit industries in China can be classified as not very concentrated industries, somewhat concentrated industries, and very concentrated industries, respectively. Meanwhile, the corresponding ratios for the United States are 10.00%, 65.00%, and 25.00%. These numbers reveal that the 4-digit manufacturing industries in China are much less concentrated across counties than those of the United States. Similar conclusions can be drawn by comparing the findings of China with those of the United Kingdom and France.

Table 2c here

In summary, using firm-level data to compile the γ_i indices, we find an increasing trend of industrial agglomeration in China for the period of 1998-2005. This is in line with and further supports the upward trend found by Bai, Du, Tao and Tong [6] in the latter half of 1985-1997 period. However, China's industrial agglomeration remains lower than those in selected developed countries. It is possible that some institutional factors such as local protectionism may interfere with the process of industrial agglomeration in China. In the next section, we shall investigate various determinants of industrial agglomeration in China.

3. Determinants of China's Industrial Agglomeration

3.1. Local protectionism versus Marshallian externalities

There are three major determinants of industrial agglomeration: Marshallian externalities, resource endowments and increasing returns to scale. In principle, there is no need for us to consider the effects of scale economies, as the index of Ellison and Glaeser measures industrial agglomeration beyond what is implied by industrial structures including the extent of scale economies. Meanwhile, the impacts of resource endowments on industrial agglomeration are not our focus either, as transportation costs are decreasing and resources become mobile across regions and countries. We thus focus on the Marshallian externalities as the key determinant of

[11] It should be pointed out, however, that our study is conducted at the firm level whereas theirs is at the plant level.

industrial agglomeration with some controls for both scale economies and resource endowments.

A fundamental condition for all of the three determinants of industrial agglomeration is the free flow of goods or services across regions. This condition is, however, not readily satisfied, as witnessed by the protests accompanying the World Bank and IMF annual meetings. In the case of China, the fiscal decentralization policy initiated in China's economic reform has given local government officials incentives for protecting local firms and industries as well as for promoting regional economic development. Indeed, as shown by Young [44], Naughton [33], and Bai, Du, Tao and Tong [6]), there was a rise of local protectionism in the earlier stage of China's economic reform, which prohibits trade across regions and impedes the process of industrial agglomeration. Au and Henderson [4] find that restrictions on rural-urban migration in China result in a surplus of labor in agriculture and lead to insufficient agglomeration of economic activity in both the rural industry and urban sectors. In what follows, we develop variables for measuring the local governments' incentives for local protectionism, variables for capturing various types of Marshallian externalities, and control variables for scale economies and resource endowments.

Local Protectionism

China was a centrally planned economy from 1949 to 1978. It operated with absolute control by China's central government over its local governments: leaders of the local governments were appointed by the central government, and were instructed to follow the plans designed by the central government. Meanwhile, most of the industrial enterprises were state-owned, and they were controlled by the central and local governments to carry out detailed work toward the central plans. All the profits of the industrial enterprises were collected by the local governments and then handed over to the central government, which then allocated budgets back to the local governments as part of the economic plan. Under this system of central planning, there was no obvious correlation between the profits collected and handed over by the local governments to the central government and the budgets they were allocated. Consequently, there was little material incentive for the local governments to pursue economic growth. Similarly, the industrial enterprises were deprived of any material incentives, and were extremely inefficient. By 1978, the Chinese economy was in a terrible state, desperate for the introduction of market incentives and economic reforms.

Since 1978, state-owned industrial firms have been able to retain some of their profits after paying various types of taxes. Non-state-owned enterprises including foreign-invested firms and China's indigenous private enterprises have been allowed to emerge and develop. Along with the enterprise reforms, there have been changes in the relations between China's central government and the local governments. While the central government still has the authority to appoint the local leaders, it needs to consult the local people. More importantly, the local governments have been able to share part of the taxes collected with the central government. The introduction of this fiscal decentralization policy has led to strong incentives for the local governments to develop regional economies on the one hand and to protect local firms and industries from regional competition on the other hand. Local protectionism impedes the free flow of goods and services across regions, and slows down the process of industrial agglomeration.

Direct measures of local protectionism are difficult to come by, as protectionist policies are often disguised in nature and tend to be qualitative rather than quantitative. Instead we look for indirect measures by focusing on the incentives of local government officials for protecting local firms and industries. An important concern for the local government officials is the degree of popular support they have from people under their jurisdictions, because the central government needs to consult the opinions of these people when appointing or promoting the local government officials. Local government expenditures, which create jobs and increase local economic activities, can win popular support for the local government officials. Thus local government officials care about how much tax revenues they can collect and how much they can keep or share with the central government.

Under China's tax system implemented since 1994, there are taxes collected and kept entirely by the local governments, including corporate income taxes from all enterprises other than those affiliated at the central government level, business tax from the sales of services, and personal income tax. There are also taxes shared between the local and central governments. The most important shared tax is the value added tax, of which 25% belongs to the local governments (Jin, Qian and Weingast [26]). The ASIF data contains four items related to taxes: value added tax, corporate income tax, business tax from the sales of services, and other taxes.[12] We define *local taxes* as the sum of (1) 25% of value added tax, (2) corporate income tax paid by local

[12] Other taxes include real estate tax, tax on vehicles, tax on land usage, and stamp tax, all of which are kept by the local governments.

state-owned enterprises, collective enterprises, and private enterprises, (3) business tax from the sales of services,[13] and (4) all other taxes. We calculate the amount of "local taxes" for each 3-digit industry, and then divide "local taxes" by total sales of the industry to illustrate the relative tax contribution of an industry to the local governments.[14]

Clearly the amount of the taxes collected and captured entirely by the local governments (items (2), (3) and (4)) for a given industry is positively correlated with the level of economic activity in the concerned industry. Meanwhile, value added value varies across industries,[15] and hence it matters whether a region has the right mix of local industries so as to maximize the shared taxes. It is expected that, in industries with higher local taxes to sales ratios, local governments have stronger incentives to protect existing local firms from competition from other regions and also have stronger incentives to nurture new local firms instead of ceding markets to firms from other regions, both of which would result in lower degrees of industrial agglomeration in those industries.

Besides favoring industries with higher "local taxes to sales" ratios, local government officials are also expected to give more protection for industries with higher percentages of state-owned enterprises. This is because in general government officials can get more private benefits from state-owned enterprises than from other types of enterprises. As shown by Shleifer and Vishny [42], it is easier for local government officials to have state-owned enterprises create more job opportunities and hire more local people than bribe private firms to do the same things. It is even more the case in China, where the official ideology of China's Communist Party places an emphasis on public ownership and its legitimacy hinges upon the continuing

[13] One important part of the business tax from the sales of services is excise on particular goods such as tobacco and liquor, which belongs to the central government. Unfortunately, the ASIF data does not provide breakdown information on excise of these special goods. Given the large amount of excise (RMB 163 billion in 2005) relative to total local taxes (RMB 953 billion in 2005), we exclude 11 3-digit industries which produce products subject to excise. These industries include tobacco processing (3-digit industry codes 161, 162, 169), alcoholic beverages (151), daily chemical products (268), jewelry processing (431), special chemical products (fireworks, 267), gasoline and diesel oil (252), tires manufacturing (291), automobile manufacturing (372), and motorcycle manufacturing (373). It is for the removal of industries subject to excise that our regression analysis is carried out at the 3-digit level.

[14] In calculating total industry sales, we exclude enterprises that belong to the central government. Central government affiliated enterprises contribute little tax to local governments because their income tax and tax from the sales of services are collected by the central government.

[15] In 2005, *value added tax to value added ratios* for 3-digit industries range from 2.5% to 22.7%, with a mean of 9.9% and a standard deviation of 3.4%. The top three 3-digit industries with the highest value added tax to value added ratios are Crude oil processing (22.7%), tobacco leaf processing (19.7%), and cement manufacturing (17.7%). The bottom three 3-digit industries with the lowest value added tax to value added ratios are meter equipment manufacturing (2.5%), ship machinery equipment (2.6%), and relay and industrial control manufacturing (2.7%).

existence and development of state-owned enterprises. Indeed, China's state-owned enterprises have been used for maintaining social stability and serving other social philanthropic purposes (Bai, Li, Tao, and Wang [7]; Bai, Lu and Tao [8]), and consequently popular support for the local government officials are higher in industries with higher degree of state ownership. We thus construct a variable *share of state-owned output* – defined as the percentage of an industry's output contributed to by state-owned enterprises – for each three-digit industry, and expect that the local government officials would like to have stronger local protectionist policies in those industries with higher degrees of state ownership, thereby hindering the process of industrial agglomeration.

<u>Marshallian Externalities</u>

Ever since the seminal work of Marshall [31], externalities have been considered as a key driver behind industrial agglomeration. Like local protectionism, however, externalities are not directly measurable except in a few exceptional cases. This leaves us to infer their existence by indirect means. There are three types of externalities: knowledge spillovers, labor market pooling, and input sharing. We shall develop variables capturing these types of externalities.

A commonly used proxy for the importance of knowledge spillovers is the proportion of R&D expenditure in total sales. However, as Feldman, Feller, Bercovitz and Burton [17] argue, formal R&D expenditure data ignore the complex process of technological accumulation and do not take into account R&D output performance. As a result, more comprehensive and outcome-based proxies of knowledge spillovers, such as innovation (Audretsch and Feldman [5]; Rosenthal and Strange [38]) are used. We use another comprehensive and outcome-based proxy of knowledge spillovers – *new products to output ratio*. In the ASIF database, output of new products is reported, and the variable "new products to output ratio" can be readily constructed. A product is identified as a new product by China's National Bureau of Statistics only if it is produced for the first time at least within a province. It is possible that some of these new products may reflect local catch-up effort in copying new products from other regions or countries. To a large extent, a significant percentage of innovation in the developing countries such as China is in essence imitation. However, this still represents a step forward in product development. We expect "*new products to output ratio"* to have a positive effect on the degree of industrial agglomeration.

Besides knowledge spillovers in R&D and product development, there are also spillovers arising from information sharing. Lovely, Rosenthal and Sharma [30] examine whether the need to acquire information contributes to spatial concentration by looking at the spatial distribution of headquarters activities. Applying differencing methods to data on headquarters and branches of exporters and non-exporters, they find that when export-related information is difficult to obtain, headquarters activities of exporters are more highly agglomerated relative to headquarters activities of non-exporters in the same industry. In the ASIF data set, the basic unit of observation is firm, not establishment, and for this reason, we cannot distinguish between headquarters and branches. Instead, we hypothesize that firms in highly export-orientated industries are more likely to concentrate, for the purpose of information sharing on foreign markets. We measure an industry's dependence on foreign markets by its *exported output to total output ratio*, and expect it to have a positive effect on the degree of industrial agglomeration.

The expected co-location of firms in export-intensive industries could also be due to the incentive to share local export infrastructures and business services, or input sharing. Indeed many of those export-intensive industries tend to have fast growth in the special economic zones designated by the Chinese government in its experiments with economic reform, and these special economic zones have easy access to transportation infrastructure out to world markets. Thus, the variable *exported output to total output ratio* may capture the externalities associated with input sharing – the second type of externalities identified by the literature.

More generally, it has been argued by Marshall [31] and others that abundant supply of specialized inputs could lead to geographic concentration of downstream firms. Holmes [23] finds a positive correlation between localization of industries and their degrees of vertical disintegration (i.e., input sharing), though the causality between agglomeration and vertical disintegration remains to be investigated. In this paper, following Holmes [23], we construct a variable called *purchased-inputs intensity* – defined as the ratio of purchased-inputs including raw materials to total output – to proxy for the degree of vertical disintegration or input sharing and expect it to have a positive effect on industrial agglomeration.

In searching for proxies for the importance of labor market pooling in an industry – the third type of externalities – one needs to identify industry characteristics that are related to the specialization of the industry's labor force. Rosenthal and Strange [38] employ three proxies: labor productivity, the percentage of management staff in the total employment, and the percentage of workers with doctorate, master, and bachelor's degrees. Unfortunately, the ASIF data does not contain any information on the education level of employees, and it does not separate employees into management staff and production workers either. To construct a proxy for the importance of labor market pooling, it is assumed that the wage level is commensurate with the skill level required in competitive industries. We define *wage premium* of an industry as the regional wage premium of an industry over the average wage in that region, averaged over all regions and weighted by the industry's employment shares in those regions. Specifically,

$$Wage\_premium_i \equiv \frac{\sum_r \frac{Wage_{ir}}{Wage_r} * Emp_{ir}}{\sum_r Emp_{ir}},$$

where $Wage_r$ and $Wage_{ir}$ are average wage of all employees in region r and average wage of employees in the 3-digit industry i in region r, respectively. Emp_{ir} is the number of employees in industry i in region r. The geographic scope is at the city level, and the industrial scope is at the 3-digit industry level. In calculating the wage premium, the wage levels at non-state-owned enterprises are used. It is because unlike state-owned enterprises (Gordon, Bai and Li [20]), non-state-owned enterprises are not shielded from market competition, and they set wage levels according to market forces. The higher the wage premium of an industry, the higher the skill level required in the concerned industry, which then implies a greater need for labor market pooling and more inclination for geographic concentration. Thus, we expect *wage premium* to have a positive effect on the degree of industrial agglomeration.

<u>Controls for scale economies and resource endowments</u>

It has been argued that geographic concentration is more significant in industries exhibiting greater scale economies. However, in principle, there is no need for us to consider the effects of scale economies, because the EG index measures industrial agglomeration beyond what is implied by industrial structures including scale economies. Indeed, for the same reason, Dumais, Ellison and Glaeser [10] and Rosenthal and Strange [38] do not include any proxy of scale economies in their empirical studies of industrial agglomeration. Recently, Alecke, Alsleben, Scharr, and Untiedt [2] have argued that the EG index is still affected by the size of an industry indirectly and in a non-linear fashion through the Herfindahl index, and they find significant effects of scale economies on industrial agglomeration even with the use of the EG index. In this paper, we construct *average firm size* – defined as the total output of an industry divided by the number of firms in the industry – as a proxy of scale economies and check if scale economies still matter with the EG index as a measure of industrial agglomeration.

It is possible that variations of *average firm size* across industries may reflect differences in capital intensity. To the extent that high capital intensity is caused by the presence of high fixed costs, which imply scale economies, then *average firm size* is a good proxy for the extent of scale economies. It is also possible that variations of *average firm size* across industries are due to the differences in the degree of state control. Indeed, despite almost thirty years of economic reform, some of China's manufacturing industries are still monopolized by a few state-owned firms, in which case the average firm size is not really a proxy for the scale economies.[16]

Variations in resource endowments across regions are traditionally considered to be an important determinant of agglomeration (Ellison and Glaeser [15]; Kim [27]). However, its underlying assumption of relatively immobile resources may be less valid than it used to be because transportation costs have declined dramatically in recent decades. For example, Glaeser and Kohlhase [19] show that costs of moving goods declined by over 90% in real terms during the twentieth century. To control for the impacts of resource endowments, we construct two proxies of natural advantages from the 1997 Input-Output table,[17] namely, *agricultural products usage ratio* and *mining products usage ratio*.[18]

Regional variations in resource endowments matter only when transportation costs are significant. Despite their importance as a control variable in studies of industrial agglomeration, transportation costs and their variations across industries in particular are not easily measured even in developed economies. Rosenthal and Strange [38] insightfully proxy transportation costs using "inventories per dollar of shipment". Presumably, industries that produce highly perishable products face high transportation costs per unit of distance, and they seek to locate close to their markets, leading to lower inventories per dollar of shipment and less industrial agglomeration.

[16] We find that average firm size of state-owned enterprises was indeed larger than that of non-state-owned enterprises, though the difference was only statistically significant in the years 2003 and 2005.

[17] China's 1997 input-output table was constructed based on flows among 124 sectors, the classification of which lies between the 2-digit and 3-digit industrial classifications. Concordance table of the 124 sectors with the 3-digit industries is used, which explains why the regression analysis is carried out at the 3-digit industry level.

[18] The agricultural category includes crop cultivation, forestry, livestock, fishery and other agriculture products; the mining category includes coal, natural gas, ferrous ore, non-ferrous ore, salt, non-mental minerals, and timber and bamboo.

While a similar variable, “finished goods to output ratio”, could be constructed from the ASIF data set, it is not to be used in this study because there are significant variations in the finished goods to output ratio between state-owned enterprises and non-state-owned enterprises. Being charged with the responsibility of maintaining social stability and burdened with surplus labor, state-owned enterprises often make products that are not really demanded by the markets and have much higher finished goods to output ratios. One could also think of using “selling expenses” - including expenses such as salespersons’ salaries and commissions, travel costs, sales office payroll and expenses, transportation costs, and advertising and promotion – as a broader proxy for the costs of reaching consumers. However, there are also issues of differences in the selling expenses between foreign-invested firms and China’s domestic firms, with the former focusing a lot more on sales if they are for China’s domestic markets. Despite the aforementioned measurement difficulties, it is a general consensus that transportation costs are decreasing rapidly in China after massive government investment in transportation infrastructures (Li [29]). Hence the roles of resource endowments are expected to diminish over time as in the case of other countries.

We summarize definitions and summary statistics of the dependent variables and independent variables in Table 3. Correlations between the dependent variable and independent variables are provided in Table 4.

Tables 3 and 4 here

3.2 Regression Analysis

Most existing studies on determinants of industrial agglomeration are based on cross-sectional data, and suffer from some difficulties of identification such as unobserved characteristics and simultaneity in data (Hanson [21]). For example, Rosenthal and Strange [38] argue that their estimation of coefficients for variables other than likely exogenous natural resource advantages and product shipping costs describe the equilibrium relationship between industry characteristics and agglomeration. In other words, some industry characteristics affect the propensity to agglomerate, while at the same time agglomeration influences these same industry characteristics.

While some of these identification problems could be dealt with panel datasets and dynamic estimation methods, it is essential for the panel datasets to have both cross-sectional and time variations. Unfortunately, our panel dataset of China's industrial firms is only over a period of eight years: 1998-2005, though one could argue that many changes did take place in these eight reform years. Thus we start by pooling all observations (firms and years), and run a simple ordinary least square regression examining the correlation between the EG index of industrial agglomeration and the independent variables constructed in Section 3.1.

The results of pooled cross-sectional regressions at the county level, city level, and province level are reported in columns 1 to 3 of Table 5 respectively. The two proxies for local protectionism, *share of state owned output* and *local taxes to sales ratio* have negative coefficients at all three geographic scopes (county level, city level and province level); and the coefficients are statistically significant in all combinations except that for *share of state owned output* at the province level. The results support the argument that local governments have stronger incentives to protect those industries with higher proportions of state ownership and those industries with more contribution to local tax revenue.

Table 5 here

Results on Marshallian externalities are largely consistent with the findings in the existing literature. At all three geographic scopes, the coefficients of *new products to output ratio* are positive and statistically significant at 1% level, supporting the important role of knowledge spillovers in industrial agglomeration. Positive and significant coefficients of *exported output to total output ratio* at all three geographic scopes strongly support the long-standing argument that the need to acquire information contributes to spatial concentration, and are consistent with the findings in Lovely, Rosenthal, and Sharma [30]. The results of *exported output to total output ratio* could also be interpreted as that industrial agglomeration is more likely in regions with better export infrastructures, supporting the role of input sharing. More generally, coefficients of *purchased-inputs intensity* are positive and statistically significant at all three geographic scopes, confirming the possibility that input sharing could be a contributing factor to localization (Holmes [23]). Lastly, the three coefficients of *wage premium* corresponding to different geographic scopes are all positive, and one of them (city level) is statistically significant. This supports the prediction that labor pooling enhances industrial agglomeration.

Results for the control variables are as expected. The coefficients of *average firm size* are positive at all three geographic scopes, and they are statistically significant at both city and province levels. These results suggest that it is still necessary to control for the scale economies even when the EG index is used as an indicator of agglomeration. Meanwhile, both *agricultural products usage ratio* and *mining products usage ratio* have positive and statistically significant coefficients at all three geographic scopes, implying that regional variations in resource endowments do matter in determining the patterns of industrial agglomeration in China's manufacturing industries. This result could be interpreted that transportation costs remain significant in China and as a result resources are relatively immobile. It could also be interpreted that local governments have protectionist policies on the sales of locally endowed resources and nurture industries that have intensive usage of these resources.

If industrial agglomeration and industrial characteristics are both endogenously affected by some common factors which have not been controlled for, then industrial characteristics are correlated with the error term and the results of the pooled cross-sectional regressions reported above are biased. One way for dealing with the potential endogeneity issue is the use of instrumental variables, which we take in the following analysis. Specifically, we use the data from China's third industrial census conducted in 1995 to construct the instrumental variables for all the independent variables discussed in Section 3.1 except those of *agricultural products usage ratio* and *mining products usage ratio*.[19] Two-stage least square estimations with those instrumental variables are then performed accordingly, and the results are shown in Columns 4 to 6 of Table 5 corresponding to the three different geographic scopes.

Our main results on local protectionism and Marshallian externalities are robust to the use of the instrumental variables. The degree of industrial agglomeration is lower for industries with higher *share of state-owned output* or higher *local taxes to sales ratios*, suggesting stronger incentives of local protectionism in these industries. Meanwhile, industries with higher *new products to output ratios* and higher *exported output to total output ratios* – proxies for knowledge spillovers – experience higher degree of agglomeration. The positive and statistically significant coefficients of *exported output to total output ratio* and *purchased-inputs intensity* also lend support to the input-sharing type of Marshallian externalities. As in the ordinary least square estimations, the effects of wage premium have been found to be generally consistent with the literature results, though the coefficients are not as statistically significant as

[19] This is because the construction of these two ratios requires the use of 1997 input-output table.

those of other proxies for Marshallian externalities. One possibility is that the proxy for labor market pooling used in this paper – wage premium – is not the most ideal as compared with those used in the literature (Rosenthal and Strange [38]).

As China has undergone significant transformations through economic reforms and integration with the world economy, it is likely that the importance of the determinants of industrial agglomeration may change over time. For example, some of the Marshallian externalities may become more prominent along with rising competitiveness of China's manufacturing and export. Meanwhile the negative impacts of local protectionism could be mitigated by the increasing competition in the market places. To explore the possibility of those changes over time, we divide our sample period into two sub-periods: 1998-2001 (first half) and 2002-2005 (second half), and re-estimate using both the ordinary least square method and the two-stage least square method with instrumental variables.

Table 6 here

The results of ordinary least square estimations are summarized in Columns 1 to 6 of Table 6 with the first three columns for the time period of 1998-2001 and the next three columns for the time period of 2002-2005. It is interesting to note that while the impacts of *share of state-owned output* and *local taxes to sales ratio* are negative in both time periods, they are statistically significant only in the first time period (i.e., 1998-2001). These results suggest that local protectionism has become less of a problem over the years. Local protectionist policies are informal tariffs for goods made in other regions, and they are only effective when the cost differences between local producers and producers of other regions are small enough. However, as China continues its economic reform and integration with the world economy, there is more entry of foreign multinationals and faster development of China's non-state-owned enterprises, both of which make it more difficult for local protectionist policies to be effective. Indeed, as presented in Section 2, the overall time trend of industrial agglomeration has been increasing over the sample period, suggesting the dominance of the forces for industrial agglomeration over those of local protectionism.

As for the proxies of Marshallian externalities, there are divergent patterns observed across the sample period. For example, *exported output to total output ratio* – a proxy for knowledge spillover on how to export to foreign markets and input sharing about export infrastructures – has consistent impacts on industrial agglomeration during the entire sample period with positive and statistically significant coefficients in both

1998-2001 and 2002-2005. The coefficients of *new products to output ratio* – a proxy for knowledge spillover – are positive in both the first half and the second half of the sample period, but they are only statistically significant in 2002-2005. One possible explanation is that knowledge spillover has become more important as China moves up the value chain and focuses on more value-added manufacturing (Rodrik, [37]; Schott, [41]) The results for the control variables (such as that for scale economies and those for resource endowments) are quite consistent across the sample period.

Columns 7 to 12 of Table 6 are the corresponding results of the two-stage least square estimations with instrumental variables constructed using data from China's third industrial census. The negative impacts of local protectionism – proxied by both *share of state-owned output* and *local taxes to sales ratio* – have become less significant over time. *Exported output to total output ratio* has had consistent impacts across the sample period, while *new products to output ratio* has gained importance increased over time. These results are very similar to those of ordinary least square estimations reported in columns 1 to 6 of Table 7, suggesting that our results are robust to controls for the potential endogeneity problems.

Taken together, the results summarized in Tables 5 and 6 show that both local protectionism and Marshallian externalities are key determinants of industrial agglomeration in China with the former becoming less significant over time whereas the latter gaining strength across the sample period.

4. Conclusions

This paper empirically examines the trends and determinants of China's industrial agglomeration using a large firm-level data set for the period of 1998-2005. We first compute the measure of agglomeration developed by Ellison and Glaeser [14]. Our results show that industrial agglomeration in China has increased consistently between 1998 and 2005. The increasing trend is robust in all combinations of industrial and geographic scopes, in contrast to the results of earlier studies such as Young [44] and Poncet [36]. Comparing with selected developed countries, however, we find that China's industrial agglomeration remains considerably lower.

Next we investigate the determinants of China's industrial agglomeration with a focus on local protectionism and Marshallian externalities. It is found that industrial agglomeration is lower in industries with greater contributions to local tax revenues and in industries with higher degrees of state ownership, suggesting the role of local protectionism among China's various regions in obstructing the process of spatial concentration of manufacturing industries. Meanwhile, there is evidence supporting the positive role of Marshallian externalities such as knowledge spillovers and input sharing in contributing to China's industrial agglomeration. Most interestingly, it is found that local protectionism has become less of a concern over time while Marshallian externalities are gaining strength across the sample period, consistent with the rising trend of China's industrial agglomeration over the sample period and possibly due to China's increasing integration with the world economy. Our results are robust to the inclusion of proxies for scale economies and resource endowments, and the use of instrumental variables for controlling the potential endogeneity problems.

Industrial agglomeration has been considered a source of sustainable competitive advantage for a national or regional economy. It has been studied extensively by economists dating back to Adam Smith. In recent years, attention has been shifted towards the political factors that may contribute or obstruct the process of industrial agglomeration. In this paper, we focus on the relation between China's central and local government that has undergone dramatic changes during China's economic reform, and investigate the role of local protectionism, unleashed by the fiscal decentralization key to China's economic reform, on the degree of China's industrial agglomeration. Our study contributes to the literature by incorporating some of the unique features in the developing and transition economy of China. It also provides evidence on the Marshallian externalities in the setting of developing economies, lending support to those governments for enacting policies nurturing the externality economies and facilitating the process of industrial agglomeration. Future work on China's industrial agglomeration should be directed at collecting data at the plant level and having a closer look at the interactions between market forces for industrial agglomeration and the political factors against it.

5. Reference

[1] Ades, A. F. and E. L. Glaeser. Trade and Circuses: Explaining Urban Giants. *Quarterly Journal of Economics* 110(1) (1995): 195-227.

[2] Alecke, B., C. Alsleben, F. Scharr and G. Untiedt. Are there really high-tech clusters? The geographic concentration of German manufacturing industries and its determinants. *Annals of Regional Science* 40(1) (2006): 19-42.

[3] Arbia, G.. The Role of Spatial effects in the Empirical Analysis of Regional Concentration. *Journal of Geographical Systems* 3(3) (2001): 271 - 281.

[4] Au, C.-C. and V. Henderson. "How Migration Restrictions Limit Agglomeration and Productivity in China." *Journal of Development Economics* **80**(2) (2006): 350-38.

[5] Audretsch, D. B. and M. P. Feldman. R&D Spillovers and the Geography of Innovation and Production. *American Economic Review* 86(3) (1996): 630-640.

[6] Bai, C.-E., Y. Du, Z.G. Tao and S. Y. Tong. Local protectionism and regional specialization: evidence from China's industries. *Journal of International Economics* 63(2) (2004): 397-417.

[7] Bai, C.-E., D. D. Li, Z.G. Tao and Y. Wang. "A Multitask Theory of State Enterprise Reform." *Journal of Comparative Economics* **28**(4) (2000): 716-738.

[8] Bai, C.-E., J. Lu and Z.G. Tao. "The Multitask Theory of State Enterprise Reform: Empirical Evidence from China." *American Economic Review* **96**(2) (2006): 353-357.

[9] Devereux, M. P., R. Griffith and H. Simpson. The geographic distribution of production activity in the UK. *Regional Science and Urban Economics* 34(5) (2004): 533-564.

[10] Dumais, G., G. Ellison and E. L. Glaeser. Geographic Concentration as a Dynamic Process. *NBER Working Paper* No. 6270, (1997)

[11] Dumais, G., G. Ellison and E. L. Glaeser. Geographic Concentration as a Dynamic Process. *Review of Economics and Statistics* 84(2) (2002): 193-204.

[12] Duranton, G. and H. G. Overman. "Testing for Localisation Using Micro-Geographic Data." *Review of Economic Studies* 72(4) (2005): 1077-1106.

[13] Duranton, G. and D. Puga. Micro-foundations of Urban Agglomeration Economies. Handbook of regional and urban economics. J. V. Henderson and J.-F. Thisse. Amsterdam, North-Holland. 4. (2004).

[14] Ellison, G. and E. L. Glaeser. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy* 105(5) (1997): 889-927.

[15] Ellison, G. and E. L. Glaeser. The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *American Economic Review* 89(2) (1999): 311-316.

[16] Fan, C. S. and X. Wei. The Law of One Price: Evidence from the Transitional Economy of China. *Review of Economics and Statistics*, **88**(4) (2006): 682-697.

[17] Feldman, M. P., I. Feller, J. E. L. Bercovitz and R. M. Burton. University-Technology Transfer and the System of Innovation. Institutions and Systems in the Geography of Innovation: Economics of Science, Technology, and Innovation. M. P. Feldman and N. Massard. Boston, MA, Kluwer Academic Publishers, (2002).

[18] Fujita, M. and D. Hu. "Regional disparity in China 1985-1994: The effects of globalization and economic liberalization." *Annals of Regional Science* **35**(1) (2001): 3-37.

[19] Glaeser, E. L. and J. E. Kohlhase. Cities, Regions and the Decline of Transport Costs. *Papers in Regional Science* 83(1) (2004): 197-228.

[20] Gordon, R. H., C.-E. Bai and D. D. Li. "Efficiency Losses from Tax Distortions vs. Government Control." *European Economics Review* **43**(4-6) (1999): 1095-1103.

[21] Hanson, G. H. Scale economies and the geographic concentration of industry. *Journal of Economic Geography* 1(3) (2001): 255-276.

[22] Holmes, T. J. The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders. *Journal of Political Economy* 106(4) (1998): 667-705.

[23] Holmes, T. J. Localization of Industry and Vertical Disintegration. *Review of Economics and Statistics* 81(2) (1999): 314-325.

[24] Holz, C. A.. China's Statistical System in Transition: Challenges, Data Problems, and Institutional Innovations. *Review of Income and Wealth* 50(3) (2004): 381-409.

[25] Holz, C. A. and Y. Lin. The 1997-1998 break in industrial statistics: Facts and appraisal. *China Economic Review* 12(4) (2001): 303-316.

[26] Jin, H., Y. Qian and B. R. Weingast. Regional decentralization and fiscal incentives: Federalism, Chinese style. *Journal of Public Economics* 89(9-10) (2005): 1719-1742.

[27] Kim, S.. Regions, resources, and economic geography: Sources of U.S. regional comparative advantage, 1880-1987. *Regional Science and Urban Economics* 29(1) (1999): 1-32.

[28] Krugman, P. Geography and Trade. Cambridge, Massachusetts, MIT Press, (1991).

[29] Li, Z.G.. Measuring the Social Return to Infrastructure Investments Using Interregional Price Gaps: a Natural Experiment. Hong Kong University Working Paper (2005).

[30] Lovely, M. E., S. S. Rosenthal and S. Sharma. Information, agglomeration, and the headquarters of U.S. exporters. *Regional Science and Urban Economics* 35(2) (2005): 167-191.

[31] Marshall, A. Principles of economics. New York, MacMillan, (1920).

[32] Maurel, F. and B. Sedillot. A measure of the geographic concentration in French manufacturing industries. *Regional Science and Urban Economics* 29(5) (1999): 575-604.

[33] Naughton, B.. How Much Can Regional Integration Do to Unify China's Markets. How Far Across the River? Chinese Policy Reform at the Millennium. N. Hope, D. Yang and M. Y. Li. Stanford, Stanford University Press**:** 204-232, (2003).

[34] Ohlin, B. Interregional and international trade. Cambridge, Harvard University Press, (1935).

[35] Pan, Z. and F. Zhang. Urban Productivity in China. *Urban Studies* 39(12) (2002): 2267-2281.

[36] Poncet, S. Measuring Chinese Domestic and International Integration. *China Economic Review* 14(1) (2003): 1-21.

[37] Rodrik, Dani, “What Is So Special about China’s Exports?” *China & World Economy,* 14(5) (2006), 1-19.

[38] Rosenthal, S. S. and W. C. Strange. The Determinants of Agglomeration. *Journal of Urban*

Economics 50(2) (2001): 191-229.

[39] Rosenthal, S. S. and W. C. Strange. Geography, industrial organization, and agglomeration. *Review of Economics and Statistics* 85(2) (2003): 377-393.

[40] Rosenthal, S. S. and W. C. Strange. Evidence on the Nature and Sources of Agglomeration Economies. Handbook of regional and urban economics. J. V. Henderson and J.-F. Thisse. Amsterdam, North-Holland. 4 (2004).

[41] Schott, Peter K. (2007), "The Relative Sophistication of Chinese Exports," *Economic Policy*, forthcoming.

[42] Shleifer, A. and R. W. Vishny. Politicians and Firms. *Quarterly Journal of Economics* **109**(4) (1994): 995-1025.

[43] Smith, A. An Inquiry into the Nature and Causes of the Wealth of Nations. London: Printed for W. Strahan, and T. Cadell, (1776).

[44] Young, A. The Razor's Edge: Distortions and Incremental Reform in the People's Republic of China. *Quarterly Journal of Economics* 115(4) (2000): 1091-1135.

Table 1a: Sample size of the data set (number of firms by year)

Number of firm \ Year	1998	1999	2000	2001	2002	2003	2004	2005
The original data including (1) mining, (2) manufacturing, and (3) production and distribution of electricity, gas, and water	164,981	161,888	162,741	171,117	181,428	196,206	270,425	265,739
The data set of manufacturing firms only	149,556	146,985	148,243	156,862	167,046	181,508	251,628	246,379
The data set of manufacturing firms only after deleting those observations with missing or zero employment	143,968	140,659	142,407	152,311	162,573	178,275	246,625	244,315
Percentage of firms with missing or zero employment	3.7%	4.3%	3.9%	2.9%	2.7%	1.8%	2.0%	0.8%

Table 1b: Number of firms by industry and year

Industry	Number of firms 1998	Number of firms 1999	Number of firms 2000	Number of firms 2001	Number of firms 2002	Number of firms 2003	Number of firms 2004	Number of firms 2005
Food Processing	11,238	10,494	9,921	9,778	9,907	10,412	13,299	13,885
Food Production	4,960	4,533	4,311	4,268	4,358	4,565	4,886	4,912
Beverage Production	3,561	3,324	3,174	3,110	3,125	3,152	3,392	3,483
Tobacco Processing	342	333	330	306	276	244	207	184
Textile Industry	10,846	10,512	10,552	11,805	12,975	14,272	22,315	21,113
Garments & Other Fiber Products	6,612	6,462	6,929	7,974	8,935	10,111	11,775	11,755
Leather, Furs, Down & Related Products	3,211	3,064	3,057	3,449	3,861	4,460	6,083	5,983
Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	2,357	2,300	2,445	2,711	2,952	3,472	4,654	5,015
Furniture Manufacturing	1,411	1,419	1,453	1,594	1,731	1,933	3,005	3,084
Papermaking & Paper Products	4,581	4,466	4,479	4,868	5,142	5,536	7,300	7,309
Printing & Record Pressing	3,754	3,646	3,551	3,562	3,687	3,780	4,564	4,400
Stationery, Educational & Sports Goods	1,751	1,762	1,837	2,003	2,290	2,584	3,676	3,603
Petroleum Processing, Coking Products, & Gas Production & Supply	1,003	955	958	994	1,113	1,279	1,932	1,930
Raw Chemical Materials & Chemical Products	10,850	10,846	10,934	11,620	12,252	13,530	18,589	18,861
Medical & Pharmaceutical Products	3,144	3,116	3,165	3,364	3,569	3,721	4,238	4,523
Chemical Fibers	782	784	807	870	897	1,001	1,733	1,543
Rubber Products	1,724	1,746	1,728	1,733	1,775	1,981	3,017	2,945
Plastic Products	5,874	5,874	6,073	6,768	7,534	8,397	12,124	12,093
Nonmetal Mineral Products	14,011	13,874	14,049	14,285	14,923	16,061	19,431	19,684
Smelting & Pressing of Ferrous Metals	3,094	2,911	2,864	3,058	3,215	3,702	6,225	5,922
Smelting & Pressing of Nonferrous Metals	2,316	2,312	2,442	2,730	2,856	3,173	4,649	4,611
Metal Products	7,919	7,870	8,119	9,097	9,832	10,946	15,885	15,734
Machinery & Equipment Manufacturing	9,076	8,884	9,092	9,841	10,579	11,899	17,981	17,737
Special Equipment Manufacturing	6,421	6,161	6,155	6,214	6,367	6,880	10,717	10,231
Transportation Equipment Manufacturing	6,519	6,378	6,565	6,948	7,322	8,045	11,532	11,200
Electric Equipment & Machinery	7,367	7,384	7,659	8,528	9,243	10,222	15,699	15,143
Electronic & Telecommunications	4,024	4,073	4,298	4,735	5,237	5,845	8,928	8,723
Instruments, Meters, Cultural & Official Machinery	1,754	1,709	1,790	1,972	2,090	2,230	3,390	3,302
Other Manufacturing	3,466	3,467	3,670	4,126	4,530	4,842	5,399	5,407
Total	**143,968**	**140,659**	**142,407**	**152,311**	**162,573**	**178,275**	**246,625**	**244,315**

Table 1c: Size of employment by year and industry

Industry	Number of employee 1998	Number of employee 1999	Number of employee 2000	Number of employee 2001	Number of employee 2002	Number of employee 2003	Number of employee 2004	Number of employee 2005
Food Processing	1,973	1,783	1,655	1,648	1,710	1,736	1,846	2,078
Food Production	1,002	955	905	891	972	1,040	1,063	1,142
Beverage Production	1,130	1,050	1,008	936	901	873	824	873
Tobacco Processing	290	278	257	246	231	211	194	195
Textile Industry	5,710	5,068	4,800	4,746	4,755	4,863	5,562	5,599
Garments & Other Fiber Products	2,112	2,024	2,153	2,366	2,648	3,005	3,271	3,458
Leather, Furs, Down & Related Products	1,103	1,093	1,125	1,265	1,408	1,647	2,043	2,221
Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	492	471	492	509	516	584	713	781
Furniture Manufacturing	249	253	270	297	338	419	652	719
Papermaking & Paper Products	1,270	1,180	1,119	1,130	1,140	1,144	1,279	1,277
Printing & Record Pressing	670	601	555	543	552	553	572	596
Stationery, Educational & Sports Goods	613	639	651	668	754	867	1,099	1,142
Petroleum Processing, Coking Products, & Gas Production & Supply	779	716	634	591	556	600	654	708
Raw Chemical Materials & Chemical Products	3,858	3,682	3,435	3,157	3,081	3,078	3,266	3,435
Medical & Pharmaceutical Products	1,025	994	990	1,020	1,045	1,073	1,062	1,154
Chemical Fibers	480	459	428	402	377	377	432	467
Rubber Products	761	707	664	612	611	626	788	779
Plastic Products	1,098	1,108	1,110	1,167	1,291	1,439	1,768	1,862
Nonmetal Mineral Products	4,502	4,310	4,083	3,889	3,860	3,916	4,081	4,151
Smelting & Pressing of Ferrous Metals	2,953	2,751	2,602	2,477	2,378	2,501	2,607	2,708
Smelting & Pressing of Nonferrous Metals	1,120	1,078	1,053	1,090	1,022	1,035	1,155	1,190
Metal Products	1,746	1,648	1,615	1,642	1,732	1,897	2,453	2,604
Machinery & Equipment Manufacturing	3,392	3,012	2,835	2,703	2,628	2,772	3,110	3,223
Special Equipment Manufacturing	2,516	2,172	2,056	1,843	1,771	1,787	2,042	2,007
Transportation Equipment Manufacturing	3,356	3,160	3,052	2,955	2,958	3,079	3,365	3,485
Electric Equipment & Machinery	2,377	2,277	2,285	2,249	2,383	2,643	3,361	3,572
Electronic & Telecommunications	1,840	1,854	1,960	2,047	2,290	2,746	3,680	4,351
Instruments, Meters, Cultural & Official Machinery	636	575	559	552	571	621	686	727
Other Manufacturing	907	916	963	1,005	1,080	1,183	1,283	1,329
Total	**49,961**	**46,813**	**45,316**	**44,646**	**45,559**	**48,315**	**54,909**	**57,833**

Note: Numbers in the table are in thousand.

Table 2a: Weighted means and summary statistics of industrial agglomeration in China's manufacturing industries (indices calculated based on employment data and weighted by employment)

Industry and Region	1998	1999	2000	2001	2002	2003	2004	2005	Change (1998-2005)
2-digit industry									
County	0.0018	0.0018	0.0022	0.0023	0.0026	0.0033	0.0046	0.0048	0.0030
City	0.0042	0.0045	0.0052	0.0058	0.0066	0.0077	0.0100	0.0104	0.0062
Province	0.0173	0.0194	0.0216	0.0230	0.0267	0.0305	0.0363	0.0370	0.0197
3-digit industry									
County	0.0043	0.0040	0.0047	0.0051	0.0059	0.0069	0.0085	0.0089	0.0046
City	0.0092	0.0094	0.0108	0.0118	0.0133	0.0151	0.0184	0.0193	0.0101
Province	0.0316	0.0341	0.0377	0.0399	0.0451	0.0500	0.0578	0.0593	0.0277
4-digit industry									
County	0.0064	0.0063	0.0077	0.0083	0.0097	0.0111	0.0131	0.0133	0.0069
City	0.0128	0.0131	0.0156	0.0172	0.0197	0.0217	0.0257	0.0263	0.0135
Province	0.0402	0.0433	0.0484	0.0510	0.0580	0.0631	0.0723	0.0741	0.0339

Table 2b: Agglomeration of China's manufacturing industries at 2-digit industry level and county level (indices calculated based on employment data)

	γ index								
Industry	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**Change (1998-2005)**
Stationery, Educational & Sports Goods	0.0194	0.0203	0.0213	0.0176	0.0175	0.0191	0.0205	0.0206	0.0012
Electronic & Telecommunications	0.0069	0.0069	0.0083	0.0095	0.0103	0.0123	0.0178	0.0175	0.0105
Leather, Furs, Down & Related Products	0.0035	0.0052	0.0058	0.0064	0.0080	0.0091	0.0105	0.0102	0.0067
Furniture Manufacturing	0.0013	0.0020	0.0019	0.0018	0.0024	0.0057	0.0076	0.0081	0.0068
Chemical Fibers	0.0008	0.0012	0.0014	0.0012	0.0023	0.0031	0.0062	0.0061	0.0053
Electric Equipment & Machinery	0.0015	0.0014	0.0018	0.0022	0.0027	0.0039	0.0052	0.0056	0.0041
Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	0.0014	0.0013	0.0017	0.0022	0.0024	0.0033	0.0053	0.0054	0.0040
Petroleum Processing, Coking Products, & Gas Production & Supply	0.0021	0.0022	0.0093	0.0045	0.0042	0.0054	0.0052	0.0051	0.0031
Food Processing	0.0010	0.0011	0.0014	0.0014	0.0019	0.0028	0.0038	0.0042	0.0032
Instruments, Meters, Cultural & Official Machinery	0.0019	0.0019	0.0022	0.0020	0.0024	0.0030	0.0037	0.0040	0.0021
Smelting & Pressing of Nonferrous Metals	0.0023	0.0024	0.0025	0.0025	0.0028	0.0030	0.0036	0.0037	0.0014
Smelting & Pressing of Ferrous Metals	0.0011	0.0011	0.0008	0.0011	0.0016	0.0022	0.0034	0.0036	0.0025
Transportation Equipment Manufacturing	0.0019	0.0019	0.0020	0.0024	0.0025	0.0026	0.0032	0.0034	0.0015
Plastic Products	0.0021	0.0027	0.0024	0.0022	0.0022	0.0021	0.0034	0.0033	0.0012
Other Manufacturing	0.0030	0.0030	0.0032	0.0033	0.0036	0.0036	0.0032	0.0032	0.0002
Nonmetal Mineral Products	0.0009	0.0009	0.0011	0.0012	0.0014	0.0018	0.0027	0.0030	0.0021
Medical & Pharmaceutical Products	0.0005	0.0003	0.0005	0.0009	0.0012	0.0016	0.0026	0.0028	0.0024
Garments & Other Fiber Products	0.0054	0.0034	0.0029	0.0027	0.0026	0.0029	0.0029	0.0028	-0.0027
Textile Industry	0.0010	0.0010	0.0010	0.0011	0.0014	0.0018	0.0027	0.0026	0.0016
Tobacco Processing	0.0001	0.0001	0.0005	0.0009	0.0010	0.0012	0.0014	0.0026	0.0025
Food Production	0.0006	0.0006	0.0008	0.0010	0.0012	0.0014	0.0020	0.0022	0.0016
Beverage Production	0.0006	0.0005	0.0006	0.0008	0.0009	0.0013	0.0022	0.0022	0.0016
Machinery & Equipment Manufacturing	0.0008	0.0008	0.0009	0.0010	0.0012	0.0015	0.0021	0.0022	0.0013
Raw Chemical Materials & Chemical Products	0.0006	0.0005	0.0009	0.0007	0.0008	0.0011	0.0020	0.0021	0.0016
Rubber Products	0.0010	0.0012	0.0015	0.0008	0.0012	0.0012	0.0019	0.0021	0.0011
Special Equipment Manufacturing	0.0009	0.0009	0.0010	0.0011	0.0013	0.0015	0.0014	0.0014	0.0006
Printing & Record Pressing	0.0000	0.0007	0.0008	0.0011	0.0015	0.0016	0.0015	0.0014	0.0014
Papermaking & Paper Products	0.0005	0.0006	0.0007	0.0011	0.0013	0.0013	0.0012	0.0013	0.0007
Metal Products	0.0007	0.0008	0.0009	0.0010	0.0012	0.0014	0.0011	0.0012	0.0005

Industry	Gini index								
	1998	1999	2000	2001	2002	2003	2004	2005	Change (2005-1998)
Stationery, Educational & Sports Goods	0.0221	0.0230	0.0236	0.0196	0.0193	0.0208	0.0219	0.0222	0.0001
Electronic & Telecommunications	0.0084	0.0083	0.0096	0.0107	0.0114	0.0135	0.0189	0.0188	0.0104
Leather, Furs, Down & Related Products	0.0050	0.0069	0.0077	0.0080	0.0094	0.0104	0.0115	0.0113	0.0063
Furniture Manufacturing	0.0037	0.0046	0.0043	0.0038	0.0042	0.0077	0.0091	0.0095	0.0059
Chemical Fibers	0.0107	0.0105	0.0118	0.0097	0.0099	0.0111	0.0120	0.0139	0.0032
Electric Equipment & Machinery	0.0024	0.0023	0.0027	0.0030	0.0035	0.0047	0.0057	0.0062	0.0038
Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	0.0037	0.0042	0.0041	0.0041	0.0039	0.0044	0.0065	0.0065	0.0028
Petroleum Processing, Coking Products, & Gas Production & Supply	0.0218	0.0231	0.0246	0.0177	0.0135	0.0127	0.0104	0.0102	-0.0116
Food Processing	0.0018	0.0017	0.0021	0.0032	0.0062	0.0038	0.0045	0.0049	0.0031
Instruments, Meters, Cultural & Official Machinery	0.0044	0.0044	0.0047	0.0042	0.0044	0.0051	0.0054	0.0058	0.0014
Smelting & Pressing of Nonferrous Metals	0.0092	0.0094	0.0099	0.0097	0.0091	0.0087	0.0083	0.0082	-0.0010
Smelting & Pressing of Ferrous Metals	0.0117	0.0124	0.0138	0.0137	0.0133	0.0118	0.0107	0.0111	-0.0006
Transportation Equipment Manufacturing	0.0045	0.0044	0.0046	0.0048	0.0046	0.0045	0.0040	0.0049	0.0004
Plastic Products	0.0028	0.0034	0.0031	0.0030	0.0028	0.0026	0.0037	0.0036	0.0008
Other Manufacturing	0.0043	0.0046	0.0046	0.0046	0.0044	0.0044	0.0041	0.0040	-0.0004
Nonmetal Mineral Products	0.0012	0.0013	0.0014	0.0015	0.0017	0.0020	0.0029	0.0032	0.0019
Medical & Pharmaceutical Products	0.0022	0.0023	0.0024	0.0029	0.0032	0.0034	0.0041	0.0044	0.0022
Garments & Other Fiber Products	0.0065	0.0040	0.0035	0.0032	0.0031	0.0033	0.0032	0.0031	-0.0034
Textile Industry	0.0014	0.0015	0.0015	0.0016	0.0019	0.0024	0.0032	0.0032	0.0018
Tobacco Processing	0.0067	0.0072	0.0073	0.0081	0.0090	0.0099	0.0130	0.0175	0.0108
Food Production	0.0017	0.0020	0.0022	0.0025	0.0030	0.0030	0.0036	0.0040	0.0023
Beverage Production	0.0020	0.0021	0.0024	0.0026	0.0031	0.0036	0.0043	0.0051	0.0031
Machinery & Equipment Manufacturing	0.0015	0.0015	0.0015	0.0017	0.0018	0.0021	0.0024	0.0025	0.0010
Raw Chemical Materials & Chemical Products	0.0018	0.0019	0.0019	0.0016	0.0017	0.0018	0.0025	0.0027	0.0009
Rubber Products	0.0039	0.0039	0.0046	0.0036	0.0042	0.0041	0.0041	0.0042	0.0003
Special Equipment Manufacturing	0.0021	0.0022	0.0028	0.0023	0.0024	0.0027	0.0022	0.0023	0.0002
Printing & Record Pressing	0.0044	0.0016	0.0018	0.0022	0.0029	0.0030	0.0026	0.0026	-0.0018
Papermaking & Paper Products	0.0014	0.0016	0.0017	0.0021	0.0025	0.0023	0.0022	0.0021	0.0007
Metal Products	0.0013	0.0013	0.0015	0.0014	0.0016	0.0017	0.0013	0.0015	0.0002

Industry	Herfindahl index								
	1998	1999	2000	2001	2002	2003	2004	2005	Change (2005-1998)
Stationery, Educational & Sports Goods	0.0027	0.0028	0.0024	0.0020	0.0018	0.0018	0.0014	0.0016	-0.0011
Electronic & Telecommunications	0.0015	0.0014	0.0013	0.0012	0.0012	0.0012	0.0011	0.0014	-0.0001
Leather, Furs, Down & Related Products	0.0015	0.0018	0.0019	0.0016	0.0014	0.0013	0.0011	0.0012	-0.0004
Furniture Manufacturing	0.0024	0.0025	0.0024	0.0020	0.0019	0.0020	0.0015	0.0015	-0.0009
Chemical Fibers	0.0099	0.0093	0.0104	0.0085	0.0076	0.0080	0.0059	0.0079	-0.0020
Electric Equipment & Machinery	0.0009	0.0008	0.0009	0.0008	0.0008	0.0007	0.0006	0.0006	-0.0002
Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	0.0023	0.0029	0.0023	0.0019	0.0015	0.0011	0.0012	0.0011	-0.0012
Petroleum Processing, Coking Products, & Gas Production & Supply	0.0198	0.0210	0.0155	0.0133	0.0094	0.0074	0.0053	0.0051	-0.0147
Food Processing	0.0007	0.0006	0.0007	0.0018	0.0043	0.0011	0.0007	0.0007	0.0000
Instruments, Meters, Cultural & Official Machinery	0.0025	0.0026	0.0025	0.0022	0.0020	0.0021	0.0017	0.0018	-0.0007
Smelting & Pressing of Nonferrous Metals	0.0069	0.0070	0.0074	0.0073	0.0063	0.0058	0.0047	0.0046	-0.0024
Smelting & Pressing of Ferrous Metals	0.0106	0.0113	0.0130	0.0126	0.0118	0.0097	0.0074	0.0075	-0.0031
Transportation Equipment Manufacturing	0.0026	0.0025	0.0026	0.0024	0.0022	0.0018	0.0008	0.0015	-0.0010
Plastic Products	0.0007	0.0007	0.0007	0.0008	0.0006	0.0006	0.0004	0.0004	-0.0003
Other Manufacturing	0.0014	0.0016	0.0014	0.0013	0.0008	0.0008	0.0008	0.0008	-0.0006
Nonmetal Mineral Products	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002	-0.0001
Medical & Pharmaceutical Products	0.0018	0.0020	0.0020	0.0020	0.0020	0.0019	0.0016	0.0016	-0.0001
Garments & Other Fiber Products	0.0011	0.0007	0.0006	0.0005	0.0005	0.0004	0.0004	0.0004	-0.0008
Textile Industry	0.0004	0.0005	0.0005	0.0005	0.0005	0.0005	0.0004	0.0006	0.0001
Tobacco Processing	0.0066	0.0071	0.0069	0.0073	0.0080	0.0087	0.0117	0.0151	0.0084
Food Production	0.0011	0.0014	0.0015	0.0016	0.0018	0.0016	0.0016	0.0018	0.0007
Beverage Production	0.0014	0.0016	0.0018	0.0019	0.0021	0.0024	0.0022	0.0030	0.0015
Machinery & Equipment Manufacturing	0.0007	0.0007	0.0007	0.0007	0.0006	0.0006	0.0004	0.0004	-0.0003
Raw Chemical Materials	0.0012	0.0014	0.0010	0.0008	0.0009	0.0007	0.0005	0.0005	-0.0007

& Chemical Products									
Rubber Products	0.0030	0.0027	0.0031	0.0028	0.0030	0.0029	0.0022	0.0022	-0.0008
Special Equipment Manufacturing	0.0012	0.0013	0.0019	0.0012	0.0012	0.0012	0.0009	0.0009	-0.0003
Printing & Record Pressing	0.0044	0.0009	0.0010	0.0011	0.0014	0.0014	0.0011	0.0012	-0.0032
Papermaking & Paper Products	0.0009	0.0010	0.0010	0.0011	0.0012	0.0010	0.0010	0.0009	0.0000
Metal Products	0.0006	0.0005	0.0005	0.0004	0.0004	0.0004	0.0003	0.0003	-0.0003

Table 2c: Comparison of γ_i index of China's manufacturing industries with those of other countries

Literature	Country	Year	Industry	Region	Percentage of industries that are		
					not very concentrated	**somewhat concentrated**	**very concentrated**
Ellison and Glaeser （1997）	U.S.	1987	459, 4-digit	3000 counties	10.00%	65.00%	25.00%
Devereux, Griffith and Simpson（2004）	U.K.	1992	211, 4-digit	477 Zip codes	65.00%	19.00%	16.00%
Maurel and Sedillot （1999）	France	1993	273, 4-digit	95 counties	50.00%	23.00%	27.00%
This paper	China	2005	537, 4-digit	2862 counties	75.98%	16.2%	7.82%

Note: Industries with $\gamma_i > 0.05$, $0.02 \le \gamma_i \le 0.05$, and $\gamma_i < 0.02$ are defined as very concentrated, somewhat concentrated, and not very concentrated, respectively.

Table 3: Definitions and summary statistics of key variables

Variable Name	Definition	N	Mean	SD	Min	Max
EG index (3-digit, county)	EG index calculated at 3-digit industry level and county level	1277	0.0083	0.0162	-0.0232	0.2798
EG index (3-digit, city)	EG index calculated at 3-digit industry level and city level	1277	0.0160	0.0240	-0.0657	0.3928
EG index (3-digit, province)	EG index calculated at 3-digit industry level and province level	1277	0.0459	0.0630	-0.4371	0.4336
Share of state owned output	(state ownership * output / sum of all types of ownership) * total output	1277	0.2741	0.2157	0	1
Local taxes to sales ratio	(25% of value added tax + corporate income tax paid by local SOEs, COEs, and private enterprises + business tax from the sales of services + all other taxes) / total sales	1277	0.0242	0.0086	0.0021	0.0825
Wage premium	$\dfrac{\sum_r \frac{Wage_{ir}}{Wage_r} * Emp_{ir}}{\sum_r Emp_{ir}}$	1277	1.0083	0.2344	0.4497	3.4842
Purchased-inputs intensity	purchased-inputs / total output	1277	0.7894	0.1168	0.4186	3.3075
New products to output ratio	total new products of an industry / total output of the industry	1277	0.0729	0.0893	0	0.6429
Exported output to total output ratio	output exported of an industry / total output of the industry	1277	0.2348	0.1984	0	0.8820
Average firm size	total output of an industry / number of firms in the industry	1277	0.3040	0.3583	0.0438	4.6567
Agricultural products usage ratio	share of inputs from agricultural sectors	1277	0.0912	0.1752	0	0.8086
Mining products usage ratio	share of inputs from mining sectors	1277	0.0531	0.1185	0	0.7180

Note: Industry variables are calculated at the 3-digit industry level.

Table 4: Correlations between dependent and independent variables

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EG Index (3-digit, county)	(1)	1									
Share of state owned output	(2)	**-0.206**	1								
Local taxes to sales ratio	(3)	**-0.092**	**0.330**	1							
Wage premium	(4)	-0.048	**0.348**	**0.074**	1						
Purchased-inputs intensity	(5)	**0.084**	**-0.134**	**-0.170**	**-0.107**	1					
New products to output ratio	(6)	0.007	**0.276**	**-0.081**	**0.392**	-0.044	1				
Exported output to total output ratio	(7)	-0.004	**0.266**	0.010	**0.216**	-0.046	**0.285**	1			
Average firm size	(8)	**0.313**	**-0.164**	**-0.209**	**-0.164**	0.050	-0.032	-0.054	1		
Agricultural products usage ratio	(9)	-0.048	-0.018	0.000	**-0.143**	-0.011	**-0.207**	**-0.084**	**-0.162**	1	
Mining products usage ratio	(10)	**0.097**	**0.183**	**0.153**	**0.071**	**0.056**	**-0.154**	0.048	**-0.267**	**-0.182**	1

Note: Coefficients in bold are significant at 5% level (2-tailed)

Table 5: Pooled cross-sectional regressions

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
	County	City	Province	County	City	Province
Share of state owned output	-0.00996 ***	-0.0103 **	-0.00836	-0.014 ***	-0.0113 **	-0.00767
	(0.0030)	(0.0043)	(0.0110)	(0.0036)	(0.0051)	(0.0130)
Local taxes to sales ratio	-0.27 ***	-0.528 ***	-0.369 *	-0.272 ***	-0.55 ***	-0.772 ***
	(0.0590)	(0.0840)	(0.2100)	(0.0750)	(0.1100)	(0.2700)
Wage premium	0.00112	0.00534 *	0.00455	-0.00224	0.00251	0.0258 **
	(0.0020)	(0.0029)	(0.0073)	(0.0030)	(0.0043)	(0.0110)
Purchased-inputs intensity	0.00938 ***	0.0222 ***	0.0354 **	0.00606	0.0204	0.0973 **
	(0.0045)	(0.0064)	(0.0160)	(0.0120)	(0.0170)	(0.0430)
New products to output ratio	0.0184 ***	0.0318 ***	0.0764 ***	0.0235 ***	0.0336 ***	0.0956 ***
	(0.0057)	(0.0081)	(0.0200)	(0.0061)	(0.0087)	(0.0220)
Exported output to	0.0328 ***	0.0605 ***	0.169 ***	0.0292 ***	0.0593 ***	0.179 ***
total output ratio	(0.0028)	(0.0040)	(0.0100)	(0.0032)	(0.0045)	(0.0110)
Average firm size	0.00226	0.0066 ***	0.0256 ***	0.00296 *	0.008 ***	0.0341 ***
	(0.0014)	(0.0020)	(0.0049)	(0.0015)	(0.0022)	(0.0055)
Agricultural products	0.00791 ***	0.0252 ***	0.0775 ***	0.00723 ***	0.0248 ***	0.0795 ***
usage ratio	(0.0026)	(0.0037)	(0.0093)	(0.0026)	(0.0037)	(0.0095)
Mining products usage ratio	0.0314 ***	0.0531 ***	0.161 ***	0.0321 ***	0.053 ***	0.162 ***
	(0.0039)	(0.0056)	(0.0140)	(0.0040)	(0.0057)	(0.0150)
year1999	0.000466	0.000133	0.0021	0.000746	0.00000997	0.00192
	(0.0017)	(0.0024)	(0.0060)	(0.0017)	(0.0024)	(0.0060)
year2000	0.000681	0.00247	0.00709	0.000335	0.00238	0.00717
	(0.0017)	(0.0024)	(0.0060)	(0.0017)	(0.0024)	(0.0061)
yr2001	0.00024	0.00141	0.00895	0.000835	0.00129	0.00978
	(0.0017)	(0.0024)	(0.0061)	(0.0017)	(0.0025)	(0.0063)
year2002	0.000414	0.00249	0.0125 **	0.000256	0.00238	0.0137 **
	(0.0017)	(0.0024)	(0.0062)	(0.0018)	(0.0025)	(0.0064)
year2003	0.000838	0.0042 *	0.0175 ***	0.0000376	0.0041	0.0196 ***
	(0.0018)	(0.0025)	(0.0063)	(0.0018)	(0.0026)	(0.0066)
year2004	0.000969	0.00474 *	0.0142 **	0.000409	0.00477	0.00136
	(0.0020)	(0.0028)	(0.0071)	(0.0029)	(0.0041)	(0.0100)
year2005	0.00336 *	0.00967 ***	0.0244 ***	0.00183	0.00919 ***	0.0248 ***
	(0.0018)	(0.0026)	(0.0066)	(0.0020)	(0.0028)	(0.0071)
Observations	1277	1277	1277	1277	1277	1277
R-squared	0.17	0.24	0.30	0.17	0.24	0.28

*, **, and *** stand for significant at 10%, 5%, and 1% respectively. Standard errors are in parentheses.
In columns (4)-(6), instruments are from the third national industrial census in 1995 for variables except for agricultural products usage ratio and mining products usage ratio. Agricultural products usage ratio and mining products usage ratio are used as instruments for themselves.

Table 6: Cross-sectional regressions on two-period average

	OLS						2SLS					
	1998-2001 average			2002-2005 average			1998-2001 average			2002-2005 average		
	County	City	Province	County	City	Province	County	City	Province	County	City	Province
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share of state owned output	-0.0154***	-0.0127*	-0.0087	-0.0005	-0.0031	-0.0034	-0.0106**	-0.0084	-0.0233	-0.0197	-0.0153	-0.0294
	(0.0049)	(0.0076)	(0.0250)	(0.0092)	(0.0130)	(0.0390)	(0.0053)	(0.0081)	(0.0260)	(0.0180)	(0.0150)	(0.0430)
Local taxes to sales ratio	-0.3000**	-0.6620***	-0.5400	-0.0523	-0.1800	-0.0141	-0.3570*	-0.7350***	-1.0890	-0.1710	-0.4010	-0.4350
	(0.1300)	(0.2000)	(0.6700)	(0.1700)	(0.2500)	(0.7000)	(0.1900)	(0.2800)	(0.7700)	(0.2000)	(0.3000)	(0.9800)
Wage premium	0.0104	0.0086	0.0203	0.0003	0.0038	0.0421	0.0029	0.0005	0.0023	0.0065	0.0009	0.0523
	(0.0048)	(0.0074)	(0.0240)	(0.0065)	(0.0095)	(0.0280)	(0.0055)	(0.0084)	(0.0270)	(0.0073)	(0.0110)	(0.0310)
Purchased-inputs intensity	0.0164	0.0035	0.0907	0.0208	0.0358	0.0984	0.0333	0.0298	0.0042	0.0155	0.0337	0.1290*
	(0.0320)	(0.0500)	(0.1600)	(0.0150)	(0.0220)	(0.0640)	(0.0670)	(0.0990)	(0.3300)	(0.0160)	(0.0240)	(0.0690)
New products to output ratio	0.0195	0.0313	0.0395	0.0301***	0.0386**	0.1090*	0.0194*	0.0244	0.0273	0.0415**	0.0552**	0.1370**
	(0.0150)	(0.0220)	(0.0550)	(0.0110)	(0.0170)	(0.0650)	(0.0120)	(0.0180)	(0.0570)	(0.0160)	(0.0240)	(0.0690)
Exported output to	0.0207***	0.0517***	0.1740***	0.0375***	0.0642***	0.1590***	0.0223***	0.0507***	0.1720***	0.0358***	0.0696***	0.1950***
total output ratio	(0.0054)	(0.0082)	(0.0270)	(0.0070)	(0.0100)	(0.0300)	(0.0061)	(0.0091)	(0.0290)	(0.0076)	(0.0110)	(0.0320)
Average firm size	0.0031	0.0055*	0.0196*	-0.0009	0.0084	0.0596***	0.0025	0.0056*	0.0255**	0.0041	0.0118*	0.0506**
	(0.0021)	(0.0032)	(0.0100)	(0.0044)	(0.0065)	(0.0190)	(0.0021)	(0.0033)	(0.0100)	(0.0047)	(0.0069)	(0.0200)
Agricultural products	0.0064	0.0226***	0.0644***	0.0101*	0.0276***	0.0884***	0.0067	0.0227***	0.0653***	0.0100	0.0301***	0.1010***
usage ratio	(0.0045)	(0.0069)	(0.0220)	(0.0060)	(0.0088)	(0.0260)	(0.0047)	(0.0071)	(0.0230)	(0.0062)	(0.0091)	(0.0270)
Mining products usage ratio	0.0343***	0.0537***	0.1440***	0.0290***	0.0522***	0.1850***	0.0346***	0.0546***	0.1400***	0.0367***	0.0624***	0.1950***
	(0.0070)	(0.0110)	(0.0350)	(0.0093)	(0.0140)	(0.0400)	(0.0082)	(0.0130)	(0.0410)	(0.0097)	(0.0140)	(0.0410)
Observations	160	160	160	160	160	160	160	160	160	160	160	160
R-squared	0.3	0.34	0.33	0.29	0.35	0.38	0.28	0.32	0.33	0.24	0.32	0.36

*, **, and *** stand for significant at 10%, 5%, and 1% respectively. Standard errors are in parentheses.

In columns (7)-(12), instruments are from the third national industrial census in 1995 for variables except for agricultural products usage ratio and mining products usage ratio. Agricultural products usage ratio and mining products usage ratio are used as instruments for themselves.