



Industrial structure and urban agglomeration: evidence from Chinese cities

Zhi Li¹ · Chengri Ding^{2,3} · Yi Niu⁴

Received: 6 July 2018 / Accepted: 22 July 2019 / Published online: 6 August 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

This paper investigates whether and how regional industrial structure/concentration influences firm productivity. Based on the firm-level data from China, the paper estimates firm productivity with regional structure conditioning on agglomeration effects and concludes that regional industrial structure plays little role in firm's output, but affects localization agglomeration, which in turn affects firm productivity. In other words, localization agglomeration is stronger in cities in which sectors are less dominated by a few large firms in their own sector. Our conclusions are robust to the classification of industries, intertemporal, and spatial dimensions of agglomeration externalities, alternative measures of regional industrial structure and agglomeration, and different spatial scales in which standard errors are clustered. An important policy implication of our findings is that China's industrial policies favoring large firms may be harmful to local economic development in the long term.

JEL Classification F1 · L1 · R1

✉ Chengri Ding
cding@umd.edu

Zhi Li
lizhiamr@163.com

Yi Niu
niu_yi_2002@163.com

- ¹ Academy of Macroeconomic Research, National Development and Reform Commission, Beijing, People's Republic of China
- ² Urban Studies and Planning Program, University of Maryland, College Park, USA
- ³ School of Public Administration and Urban Development Research Center, Zhengzhou University, Zhengzhou, People's Republic of China
- ⁴ International School of Economics and Management, Capital University of Economics and Business, Beijing, People's Republic of China

1 Introduction

Regional industrial structure is, in theory, of great importance to firm productivity and economic growth potential. In a competitive industrial structure that is an alternative to oligopoly and exhibits a high degree of vertical disintegration and specialization, agglomeration economies are constructive for small firms to compensate for their disadvantages in scale economies (Carree and Thurik 1999). A diverse industrial structure would also enhance productivity and promote innovation through cross-fertilization, have a positive impact on value-added growth, and be very important for the development of high-tech industry and the attraction of new firms (Jacobs 1969; Henderson et al. 1995; Duranton and Puga 2001; Batisse 2002; Puga 2010). In contrast, regional industrial concentration may negatively affect firm productivity and local economic growth (Chinitz 1961; Rosenthal and Strange 2003). A region dominated by a few large firms in an industry would generate fewer positive externalities associated with agglomeration economies, ultimately diminishing productivity and obstructing entrepreneurship and innovation (Chinitz 1961; Saxenian 1994). This implies that the regional industrial structure may help to explain differences in economic performance and dynamics between regions with a similar level of industrial agglomeration (Chinitz 1961; Saxenian 1994; Rantisi 2002).

Theoretically, a highly concentrated industrial structure may weaken agglomeration economies by limiting the degree of intermediate input sharing, labor market pooling, and knowledge spillover. The reasons are following. First, large firms have economic advantages in seeking nonlocal suppliers for their intermediate inputs, which then reduces the size of the local market of independent suppliers (Enright 1995; Porter 1998). On the other hand, large manufacturers may offer long-term and large-volume contract on supply and therefore are more attractive to local suppliers than small manufacturers (Booth 1986). Local suppliers are more responsive to the input demand of large manufacturers than small ones, and the presence of large manufacturers may raise input costs for small ones in the same region (Lee et al. 2010).

Second, dominant firms in a region may negatively affect labor-pooling economies for small firms in the same industry. This is because workers, especially those with skills and experience, are more attracted to large firms that offer better employment compensation packages and more stable employment opportunities than small firms (Booth 1986; Audretsch 2001). Third, a regional economy dominated by large firms tends to reduce knowledge spillovers and nurse business cultures that discourage innovation and adjustment to changing markets (Saxenian 1994; Chinitz 1961; Carree and Thurik 1999). Large dominant firms tend to be vertically integrated, which decreases face-to-face contact across firms (Enright 1995). If an industry is dominated by a few large inward-looking firms in a region, all other firms may suffer from lack of flexibility and be insensitive to innovation (Porter 1998). A concentrated industrial structure would also limit entrepreneurship (Chinitz 1961), which

further retards the generation of new products and technologies as small firms are major sources of innovation and industrial evolution (Acs 1992; Audretsch 2001).¹

Empirically, the literature produces a few papers with limited and mixed conclusions on the impact of industrial structure on firm productivity. Feser (2002) identifies a significantly negative relationship between industrial dominance and plant-level productivity for the innovative-intensive measuring and controlling devices industry but no significant influence of industrial concentration for the low-tech farm and garden machinery industry. Drucker and Feser (2012) and Drucker (2013) conduct cross-section regressions that examine both the direct effect of industrial dominance on plant-level productivity and its intervening effect through limiting agglomeration economies. They conclude that a concentrated regional industrial structure is directly associated with lower productivity, but does not limit the types of agglomeration economies measured, such as input sharing, labor pooling, and knowledge spillover. Carree and Thurik (1999) reveal that industrial concentration matters for national industrial output, but the direction of the impact differs between developed and developing countries. Gopinath et al. (2004) conclude a nonlinear relationship between industrial concentration and productivity growth in US manufacturing industries.

This paper investigates the impact of regional industrial structure on firm productivity. More specifically, the paper attempts to answer the following research questions: (1) Does regional industrial structure affect firm productivity? (2) Does regional industrial structure affect localization and urbanization economies? and (3) Does the impact of industrial structure on firm productivity vary across types of firms or periods?

Those questions will be examined by using the firm-level data of China. China offers a unique case to examine these questions. The reasons are threefold. First, the local industrial dominance underwent dramatic changes during the market-oriented reform, but its effects have not been thoroughly investigated. Important causes of the changes include, but not limited to, the privatization of state-owned enterprises, integration to global markets (e.g., entering WTO), and the gradual removal of government protection. In general, the local own-industry dominance became weaker, as the indicator, a Herfindahl index calculated in the research, shows a median value of 0.13 in 1998 and 0.07 in 2007. Second, China has adopted strong industrial policies since 1978 to promote its economic growth. This is manifested by the designation of numerous special development zones throughout the country such as industrial parks, high-tech districts, export-oriented zones, and so forth. A regional dominant firm thus may get preferential treatment through taxes and loans, and legal protection from local governments. Third, local industrial policies in less-developed areas in China mainly focus on attracting large firms. Although these strategies may promote economic growth in a short run, the large and often dominant firms in these areas are always found to be resource-oriented and vertically integrated and hardly contribute to local industrial innovation or entrepreneurial activities (Gai et al. 2015;

¹ One may wonder whether the presence of dominant firms discourages other local firms' intention in the improvement of productivity. While this may be possible, this paper does not find the evidence that the dominance directly lowers firm productivity, but suggests that the dominance reduces agglomeration economies.

Tang 2011).² The negative influence of dominant firms, however, remains underexplored in the literature on agglomeration economies in China. On the other hand, our findings will have general implications given the increasingly dominant private sector and a well-established market economy. In the industrial sector, for instance, state-owned and state-controlled enterprises contributed 34% of total output in 1995, but only 12% in 2010 and 8% in 2017. The private economy provided more than 80% of urban jobs and more than 60% of GDP in 2017.

This paper is organized as follows. The next section discusses the literature; Sect. 3 presents data, measures, and models; Sect. 4 discusses results, and finally policy implications and conclusions are presented in Sect. 5.

2 The literature

Chinitz (1961) is among the first studies that pointed out the importance of regional industrial dominance to the growth potential of firms. He argues that New York is much more entrepreneurial than Pittsburgh as the former is distinguished by a large number of small firms and the latter is dominated by a few large integrated steel firms. The concentrated industrial structure in Pittsburgh limits the availability of capital for startups, reduces input sharing and labor pooling, and impedes entrepreneurial activities. In contrast, Rantisi (2002) reports that more than 85% of firms in New York's garment district are small- to medium-sized enterprises, which compete with each other, benefit from specialized labor and services, and monitor rival firms' performance and practices.

Saxenian (1994) compares Boston's Route 128 and northern California's Silicon Valley and concludes that the fundamental difference in the regional industrial structure is the key factor to explain their diverging growth trajectories in the 1980s. Both areas were centers of electronics and high-tech industries. When facing the changing market, Silicon Valley successfully made the transition to software and other computer-related industries, generating a large number of successful startups. In contrast, Route 128 failed to make the transition to smaller workstations and personal computers, which led to continuous stagnation and decline. Silicon Valley has a regional network-based industrial structure that promotes collective learning and horizontal communication among different firms. The Route 128 region is dominated by a few large corporations that perform their own work, and the industrial structure is relatively rigid and hierarchical. Thus, it is difficult for firms to adjust to the changing market and for small firms to survive in the Route 128 area.

² In December 2018, we conducted a field survey for this research in Gansu Province, an underdeveloped province in China, and found that most local economies are dominated by one or few large firms and have been suffering economic stagnation in recent years. For example, according to the information provided by the bureau of finance of Honggu District in Lanzhou City, the capital city of Gansu Province, the largest firm, Fangda Carbon New Material Technology, contributes about 62% of the total local tax revenue in 2017. However, this firm is the only one in its industrial group and has little connection with other local firms or industries. Both the supplier and consumer industries are located in other regions.

Quantitative studies on the impact of regional industrial structure on agglomeration economies yielded mixed and inconclusive results. Henderson (2003) shows that agglomeration economies are mainly created through the large number of establishments rather than their large size. Compared to medium-sized or large establishments, Rosenthal and Strange (2003) argue that small firms generate greater external effects by finding that the concentration of own-industry employment in small establishments has a larger positive impact on both the birth and employment of new firms. Carree and Thurik (1999) reveal that the influence of industrial concentration (employment share of large firms) on national industry output can be either positive or negative in twelve European countries, depending on the industries. They conclude that industrial concentration may enhance production in less-developed countries and impede production in more developed countries. Gopinath et al. (2004) demonstrate an invert U-shaped relationship between industrial concentration, which is measured by a four-firm concentration ratio, and productivity growth in US manufacturing industries. Acs and Audretsch (1988) report a negative relationship between the concentration and innovation rate at four-digit industry level. Levin et al. (1985) and Lee (2005) exhibit that the effects of industrial concentration on innovation vary with industry conditions, cost of imitation, patent protection, and the importance of a firm's technological competence in an industry. By measuring local competition as the number of firms per worker, studies suggest that a more competitive industrial structure would promote both the employment growth (Glaeser et al. 1992) and firm birth (Rosenthal and Strange 2003; Glaeser and Kerr 2009).³

A few empirical studies have directly examined the relationship between regional industrial dominance and the performance of firms in the same industry. Feser (2002), Drucker and Feser (2012), and Drucker (2013) are the studies directly incorporating both regional industrial dominance and agglomeration economies into the plant-level production function. Feser (2002) measures industrial concentration as the share of the total sales made by the four largest firms (four-firm concentration ratio) in a commuting zone and uses plant-level data in 1992 to investigate two industries in the USA—the farm and garden machinery industry and the measuring and controlling devices industry. Feser finds a significantly negative relationship between regional industrial dominance and plant-level productivity for the innovation-intensive measuring and controlling devices industry but no significant influence of industrial concentration for the low-tech farm and garden machinery industry. Drucker and Feser (2012) and Drucker (2013) use plant-level data in 1992, 1997, and 2002 for three sectors in the U.S.—rubber and plastics, metalworking machinery, and measuring and controlling devices. According to the firms' shipment values, the researchers measure the concentration by three indicators; the five largest firm concentration ratio, the Herfindahl index, and the Rosenbluth index. Based on cross-section regressions by year, they conclude that a concentrated regional

³ A large part of these quantitative researches used aggregate data instead of micro-level data. It might be one reason for the mixed results, because using micro-level data helps to improve the measure of industrial structure and reduce aggregation bias in the estimation.

industrial structure is directly associated with lower productivity, but that it does not limit the three sources of agglomeration economies that they are able to measure, including input sharing, labor pooling and knowledge spillover.

There are plenty of studies on industrial agglomeration in China. Industries in China exhibit agglomeration economies, but the magnitude of their effects appears to be weaker than in developed countries like France, the UK, and the USA. The reason may be that Chinese industries are subject to high degree of market fragmentation and regional protectionism (Lu and Tao 2009; Long and Zhang 2012). Ke (2010) reveals that industrial agglomeration and the city-level labor productivity are positively, mutually, and causally related, which implies that industries tend to concentrate toward more productive areas to achieve higher productivity. Studies examining agglomeration effects on firm performance in China have found that agglomeration could promote or discourage productivity, depending on the size or types of agglomeration (Batisse 2002; Fu and Hong 2011; Lin et al. 2011; Yang et al. 2013; Hu et al. 2015). Chen and Wu (2014) suggest agglomeration may help to achieve social goals, because they find evidence that agglomeration in China increases a firm's pension contribution. We do not find any studies on the relationship between regional industrial structure, agglomeration economies, and firm productivity in China. Our research fills in the gap.

3 Data, variables, and model specification

3.1 Data

The data used in this study come from the China Annual Survey of Industrial Firms collected by the National Bureau of Statistics (NBS). The data contain all the state-owned enterprises (SOEs) and those non-state-owned enterprises (non-SOEs) with annual sales greater than 5 million RMB, the so-called above-scale firms.⁴ It reports detailed information for individual firms, including location information, industrial code, input and output of production, etc. This firm-level data serves as the cornerstone of many aggregate industry data used and published by NBS, and comparing our data with NBS's aggregate data for many key variables indicates a high level of consistency. Holz (2008) considers the data as quite reliable, particularly after 1998, because it includes large enterprises with a well-established accounting system, and these enterprises report regularly to the statistical authority so that inconsistencies can easily be detected. Since 2007, these data have been used extensively in many publications and earns trust among Chinese and international scholars (Brandt et al. 2014; Ding and Niu 2019).

⁴ A comparison to the First Economic Census indicates that, in 2004, above-scale firms represent 22% of all registered firms, accounting for 78% of employment, and produce 90% of output in industrial sectors. As 53% GDP was produced by the secondary sector in 2004, firms in our data represent a crucial component in China's economy.

Manufacturing firms in the period between 2000 and 2007 are used. We do not use data after 2007, because many key variables, such as value-added and material inputs, are no longer available. We exclude the data in 1998 and 1999, because the shares of SOEs are quite large in those 2 years, exceeding 30%. The 2 years, however, are used in constructing variables of lagged agglomeration. We drop industries in the mining, electricity, gas and water sectors, because the location and productivity of mining firms largely depend on natural conditions, and firms in the electricity, gas and water sector are mostly government owned.

While different researches do not use exactly the same methods in cleaning and processing the data, we primarily follow the steps of Brandt et al. (2014). It mainly involves four types of work. First, we link firms over time by using not only the unique ID assigned by NBS, but also the combination of other information such as firm's name, phone number, address, etc. Second, nominal variables of output and input (raw materials and intermediate inputs) are deflated by very refined industries. Third, much external data, including the 1993 annual enterprise survey data, the nominal capital stock at the two-digit industry level by province are combined with the firms' nominal capital stock to estimate the real capital stock of firms.⁵ Finally, we fix data problems or errors including typos, missing values, changes of industrial classification in 2002, and changes of geographic code that is used to identify the locations of firms.

3.2 Variables

3.2.1 Measuring regional industrial dominance

We used the Herfindahl index of sales in a city to measure regional industrial concentration. The index in city j is given by

$$\text{Herfindahl index}_j = \sum_{i \in j} \left(\frac{\text{Sale}_i}{\sum_{i \in j} \text{Sale}_i} \right)^2,$$

where Sale_i is the sales of firm i located in city j , and $\sum_{i \in j} \text{Sale}_i$ denotes the total sales of city j .

We chose the Herfindahl index mainly because it considers the full distribution of firm size and is superior to the concentration ratio in reflecting market structure, as suggested by Scherer and Ross (1990) and Amato (1995). As the firms in our data represent most of the sales (90.9% in 2004), the influence of missing small firms should be trivial, if any.⁶ In the comparison of several indicators for industrial structure, Amato (1995) suggests that the leading firm share is sometimes also a good

⁵ See Section 3.3 of Brandt et al. (2014) for the details in the measurement of the real capital stock as well as the inspection on the measurement.

⁶ In the Herfindahl index, we use sales instead of other variables like employment for following reasons. First, sales are probably a direct measure on market dominance than employment. Second, using sales

measurement. To check the robustness of our results to alternative indicators, we use the leading firm's share, defined as the share of the largest firm in a city.

3.2.2 Measuring agglomeration economies

The spatial scope of measuring agglomeration in this paper is city. We define the city by prefecture's city proper and county-level cities. The number of cities varies between 650 and 670 in the study period. These cities contain approximately three-fourths of manufacturing firms in our data. In 2000, they occupied 18% of land area, and produced 84% of GDP in industrial and service sectors. We exclude counties that are mainly rural areas. Doing so, as shown later, we capture most of the effects of agglomeration economies in our estimation. By the China's industrial classification, there are 30 2-digit manufacturing sectors and 162 3-digit sectors. We focus on 20 digit industrial sectors and use 3-digit industries for robustness check.

We use total employment in own-industry activities as a proxy to localization agglomeration and total employment in other-industry activities as a proxy to urbanization agglomeration. Alternatively, agglomeration can be proxied by the number of firms and population. We use employment because our data includes the majority of manufacturing workers but only a fraction of manufacturing firms.⁷ Besides agglomeration economies, what agglomeration generates may also involve negative consequences such as congestion. So what we estimate will be the net effect from agglomeration.

Employment size is probably a better measure of urban agglomeration in China than employment density. The boundaries of Chinese cities (city proper and county-level cities) are administratively delineated. A city typically consists of an urban core or built-up areas, surrounded by rural areas (Ding 2013). The density of the city, therefore, is heavily affected by the size of rural areas. This problem has been severe as many adjacent rural areas (counties or towns) are administratively merged into city proper in the past decades. Merging and/or annexing a county could increase the area of a city proper by more than one-third and dramatically reduce its density without any significant changes to its population or employment. Density is not a good measure for agglomeration in China also because the government-led land development often results in excessive land supply (Ding 2007).

Footnote 6 (continued)

is consistent with the literature, so that our results are comparable to others (Drucker and Feser 2012; Drucker 2013). Third, firms in our data represent nearly 80% of total employment but more than 90% of total sales, so using sales improves the measurement in this paper.

⁷ The quality of employment data is better than population data (from China City Statistics Yearbooks or China City Construction Statistics Yearbooks). The city population data is notoriously inaccurate as they count only the registered population (residents with city *hukou* or temporary residents living at city over 6 months in a year) (Ding and Li 2019). Massive floating population is excluded from city statistics data. Therefore, using population data is likely to underestimate the agglomeration of large cities that attract many migrant workers and overestimate agglomeration of small cities that are losing population and labor force.

Table 1 Descriptive statistics of variables

	Obs.	Mean	SD	Min	Max
Value added, thousand of RMB	19,55,669	22,431.12	2,09,203.70	0	5.92E+07
Log (value added)	19,43,333	8.6190	1.4981	-7.0957	17.8963
Employment	22,20,014	282.76	1324.03	0	1,94,410
Log (employment)	21,99,340	4.7648	1.1785	0	12.1777
Capital, thousand of RMB	21,92,700	39,268.75	5,77,141.30	0.4155	1.30E+08
Log (capital)	21,92,700	8.5248	1.7054	-0.8783	18.6815
Herfindahl index on sales in own industry	15,46,323	0.1757	0.2155	0.0033	1
Herfindahl index on sales in other industries	16,25,639	0.0569	0.0940	0.0018	1
Localization agglomeration: log (employment in own industry)	15,67,758	9.0920	2.2040	0	13.9907
Urbanization agglomeration: log (employment in other manufacturing industries)	15,47,653	12.0973	1.5866	2.3979	14.9195

Herfindahl index, localization agglomeration and urbanization agglomeration have fewer observations because firms outside cities are excluded. We compute localization agglomeration for a firm by $\log(\text{own industry employment in the same city} - \text{the firm's employment} + 1)$. So, for a firm with no other own-industry firms in the same city, its localization equals zero. Similarly, urbanization agglomeration for a firm equals $\log(\text{total employment in other manufacturing industries in the same city} + 1)$

Table 1 provides descriptive statistics of key variables. It illustrates that firm size varies substantially, as indicated by very large standard deviations relative to the mean values of value added, employment, and capital. At the mean level, a firm hires 283 workers, uses capital that worth approximately 40 thousand RMB, and produces value added slightly above 22 thousand RMB. In the same city, the firm has, on average, nearly nine thousand own-industry workers, and almost 179 thousand other-industry workers. The average Herfindahl index is about 0.18 in own industry and about 0.06 in other industries. The variables of Herfindahl index and agglomeration contain large variation, as shown by the standard deviation.

3.3 Empirical specification

A direct way to test the effects of external environment such as agglomeration economies on firm productivity is to estimate the production function (Rosenthal and Strange 2004). To make our empirical results comparable to the literature, similar to Henderson (2003), Drucker and Feser (2012) and Drucker (2013), we estimate the following model where the dependent variable is a firm's output, and independent variables include inputs as well as measures of agglomeration economies and industrial concentration:

$$\ln(VA_{it}) = \beta_1 H_{jt} + \beta_2 A_{jt} + \beta_3 (H_{jt} * A_{jt}) + \gamma \ln(X_{it}) + \delta_t + \theta_{ij} + \varepsilon_{ijt} \quad (1)$$

for firm i in location j at time t . In function (1), the dependent variable VA_{it} represents the firm's value added. The variable H_{jt} refers to the Herfindahl index, and A_{jt} refers to measures of agglomeration. We use the interactive term $H_{jt} * A_{jt}$ to examine how regional industry dominance affects the productivity effect of agglomeration economies. As we have both localization and urbanization agglomeration (L_{jt} and H_{jt}), we have two interaction terms of $H_{jt} * L_{jt}$ and $H_{jt} * U_{jt}$. The term X_{it} includes a firm's input data such as number of workers and capital. The terms of δ_t , θ_{ij} and ε_{ijt} capture those unobservables that influence the firm's output. Variable δ_t varies only across periods, θ_{ij} varies only across firms, and ε_{ijt} varies across both firms and periods.

The specification of (1) implicitly assumes Hicks' neutrality, i.e., the changes in regional industrial environment do not affect the firm's input choice. To test the validity of this assumption, we let localization agglomeration and urbanization agglomeration interact with a firm's inputs X_{it} in estimating function (1). Experimental results suggest that these interaction terms all have very small coefficients (between -0.003 and 0.005) and are statistically insignificant. We hence believe that our assumption of Hick's neutrality is reasonable.

Endogeneity may be an issue in the estimation of function (1). First, several unobserved time-invariant factors may influence regional industrial environment and firm productivity simultaneously. For example, high-ability entrepreneurs may not locate randomly but prefer locations with certain features such as a mild climate and business-friendly environment. On the other hand, local amenities, local resources, or business culture could help to foster large dominant firms and affect

firm productivity. To control the firm fixed effects (θ_{ij}), we use the within estimator for panel data.⁸

China's vast urbanization and rapid economic growth generate considerable time variation for key variables, which is needed to validate the fixed-effects estimator. During our study period of 2000–2007, the urbanization rate of China increased by 9.7% (from 36.2 to 45.9%), and urban population rose by 32.1% (from 459 to 606 million). The national GDP and manufacturing GDP grew by 1.69 times and 1.77 times, respectively, and GDP per capita also increased by 1.58 times, during the period. This general pattern is well echoed by our microdata of firms. From 2000 to 2007, the median value of own-industry employment in the same city expanded by 83.3%, and the median value of other-industry employment expanded by 56.2%. The industrial concentration also dramatically changed. Consider the own-industry Herfindahl index. From 2000 to 2007, its 25th, 50th and 75th percentiles declined from 0.039 to 0.023, from 0.117 to 0.066, and from 0.282 to 0.161, respectively. Similarly, significant changes also appear in the other-industry Herfindahl index, whose 25th, 50th and 75th percentiles declined from 0.0140 to 0.0096, from 0.0315 to 0.0190, and from 0.0718 to 0.0444.

Second, unobserved time-variant factors may lead to the endogeneity. The market fragmentation at the province level in China is likely to generate differences in prices of output and intermediate inputs (Young 2000). Such price differences not only act as shocks to value added VA_{it} measured in monetary terms, but also affects firm's input choice X_{it} as well as the regional industrial environment H_{jt} and A_{jt} . The development of regional transportation affects firm productivity and regional industrial environment. To control for these factors, we add province-year fixed effects.⁹ In addition, we also use industry-year fixed effects to control for national shocks to productivity and allow these shocks to differ across 2-digit industries. After controlling for firm/location fixed effects, province-year fixed effects, and industry-year fixed effects, what ε_{ijt} contains is contemporaneous idiosyncratic firm output shocks.

Even though the estimation with fixed effects will address the endogeneity issue, we still try to apply instrumental variables to test the robustness of our results. It is noted that finding valid instruments for regional industrial environment and a firm's inputs is difficult (Henderson 2003; Puga 2010). We chose lagged variables as instruments and use the first difference estimation to control for firm/location fixed effects. So the variables in function (1) become the difference between periods t and $t - 1$, and instruments are the predetermined variable at $t - 2$. By doing so, we assume that predetermined values (at $t - 2$) of regional industrial environment and input choices

⁸ Certainly, in the decision of using the fixed effects estimator, there is a tradeoff between preserving the variation versus controlling unobserved factors. As the unobserved factors are likely to significantly influence firm productivity and the local environment (industrial concentration, agglomeration, etc.) at the same time, we chose to control fixed effects. This choice is also consistent with many influential works, for example, Henderson (2003), as discussed by Rosenthal and Strange (2004). In the following paragraph, we show the within variation in our data should be large enough to identify the key parameters.

⁹ After excluding three provinces of Qinghai, Xinjiang and Tibet that have extremely low population density, we end up with 29 provinces or equivalents, as compared to 650–670 cities.

influence their future changes (from $t - 1$ to t), but these predetermined values are exogenous to the residual.¹⁰ Results of GMM estimator, however, suggest an issue of weak instrument that may further increase bias in estimated coefficients and enlarge asymptotic standard errors (Staiger and Stock 1997). Nevertheless, the signs of key variables in GMM estimates are the same with those in fixed-effects estimates. We thus focus on our discussions on the results with fixed effects.

Whether or not the residual is correlated across firms or periods is crucial for the consistent estimation of standard errors of coefficients. Given the fixed effects that we have controlled, the intertemporal correlation of the residual for a given firm should be relatively weak. Since the key variables H_{jt} and A_{jt} are computed at the city level, overlooking the correlation of the residual between firms in the same city may significantly underestimate the standard errors even if the correlation is not strong (Moulton 1986). Thus we report standard errors clustered at the city-year level in our main results. We also let the standard errors cluster at the city level, province-year level, or province level to check the robustness of our results.

4 Results

4.1 Basic results

We first report estimated results for function (1) in Table 2. The estimated results reveal the following findings.

First, industrial concentration in the own industry has negative effects on firm productivity. In columns (1)–(2) that only include the own-industry concentration, we find it negatively correlated with value added. This negative correlation results from an indirect mechanism that the concentration retards localization economies, as indicated by the negative and statistically significant coefficients for the interactive term between the concentration and localization in columns (5)–(8). The magnitude of the concentration's effect on localization economies is large. The results in column (6) suggest that a one standard deviation increase in the own-industry concentration will lead to a reduction of approximately 40% on the elasticity of localization economies. The coefficients for the concentration across columns (5)–(8) all have very low significance, implying that the dominance does not directly affect firm productivity.

Second, industrial concentration in other manufacturing industries has little effect on firm productivity. Column (7) adds other-industry concentration, urbanization agglomeration, and their interaction term into independent variables, and controls for all three types of fixed effects. Our results reveal that neither other-industry

¹⁰ Using longer lags may better satisfy the exclusion restriction, but there are at least two considerations in this research. The first is the whether the instrument is strongly correlated with the endogenous variable. Using longer lags generally makes the correlation weaker. The second consideration is data. This research uses a 10-year panel data, so the longest lag is 8 years in the first difference estimation. Further, when using longer lags, we have to drop more years of samples.

Table 2 Effects of industrial dominance on firm productivity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Dependent variable: log (value added)</i>									
Dominance in own industry: Herfindahl index on sales	-0.6926 (0.0225)**	-0.2186 (0.0146)**			0.0549 (0.0452)	0.0316 (0.0299)	0.0288 (0.0294)	0.0416 (0.0319)	
Localization: Log (employment in own industry)			0.0337 (0.0022)**	0.0119 (0.0016)**	0.0539 (0.0053)**	0.0252 (0.0036)**	0.0242 (0.0035)**	0.0271 (0.0041)**	
Interaction between localization and own-industry dominance					-0.0747 (0.0057)**	-0.0362 (0.0039)**	-0.0359 (0.0038)**	-0.0379 (0.0043)**	
Dominance in other manufacturing industries: Herfindahl index on sales			0.5658 (0.0215)**	0.0414 (0.0159)**	0.5322 (0.0209)**	0.0302 (0.0039)**	0.0284 (0.0167)	0.0439 (0.0175)*	
Urbanization: Log (employment in other manufacturing industries)							-0.0143 (0.0377)		-0.0228 (0.03837)
Interaction between urbanization and other-industry dominance									
Log (employment)	0.5031 (0.0060)**	0.5139 (0.0057)**	0.4826 (0.0057)**	0.5128 (0.0057)**	0.4830 (0.0057)**	0.5123 (0.0115)**	0.5124 (0.0057)**	0.5129 (0.0057)**	0.5132 (0.0057)**
Log (capital)	0.3513 (0.0046)**	0.1996 (0.0029)**	0.3006 (0.0044)**	0.1996 (0.0029)**	0.2976 (0.0043)**	0.1993 (0.0053)**	0.1992 (0.0029)**	0.1993 (0.0029)**	0.1996 (0.0029)**
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year fixed effects	No	Yes	No	Yes	No	Yes	Yes	Yes	Yes
Province-year fixed effects	No	Yes	No	Yes	No	Yes	Yes	Yes	Yes
R ²	0.8205	0.8502	0.8255	0.8501	0.8261	0.8502	0.8502	0.8502	0.8501
Obs.	11,71,807	11,71,807	11,71,807	11,71,807	11,71,807	11,71,808	11,71,787	11,71,808	11,71,787

Observations are individual firms. Numbers in parentheses are robust standard errors clustered at the city-year level

**Statistical significance at 1% level, and *significance at 5% level

concentration nor the interaction term has a statistically significant coefficient. Their significance remains low in column (9) where we drop the variables of own-industry concentration, localization agglomeration, and their interaction.¹¹

Third, we find evidence for both localization economies and urbanization economies. According to the results in column (6), we find that at the mean level, the elasticity of localization agglomeration to value added is approximately 2%, and the elasticity of urbanization agglomeration is approximately 3%. We also find that missing industry-year and province-year fixed effects bring unbelievably high coefficient for urbanization economies, and in results not reported here, missing firm fixed effects or the interaction term between the concentration and localization economies may severely underestimate agglomeration economies. Perhaps due to the collinearity when other-industry concentration and the interaction term are present, the urbanization agglomeration variable loses some significance in column (7), but is still significant at the 10% level.

The inclusion of fixed effects greatly contributes to the identification of our empirical model. We find that, in results not reported here, only 45% of the dependent variable's variation is explained by own-industry concentration and firm's inputs, and then adding industry-year fixed effects and province-year fixed effects increases the *R* square by about 10%. After adding firm/location fixed effects, *R* square rises by another 30%. It implies that the unobserved heterogeneity of firms and/or locations probably have important influences on firm output, the concentration, and agglomeration. In total, implementing all three types of fixed effects nearly doubles the explanatory power of our empirical model (*R* square from 45 to 85%). Those results imply that overlooking these unobserved factors, particularly the firm/location fixed effects, is likely to generate biased estimates of coefficients. On the other hand, since we use the within variation to identify the parameters, the estimated effects of industrial concentration are perhaps the short-term effects. In that case, we might underestimate the effects of industrial concentration, if the long-term effects are present but not properly identified by the within estimator. The high statistical significance of key variables given so many fixed effects also makes our empirical findings more convincing.

The coefficients of employment and capital inputs reveal the following findings. First, all of the coefficients are positive and significant, consistent with our expectations. Second, the sum of the two coefficients is less than one, implying decreasing returns to scale. Third, the coefficients change significantly after controlling fixed effects. The coefficient of capital declines by approximately 43% from column (1) to column (7). An interpretation is that unobserved variables, such as entrepreneurial ability and local accessibility, may simultaneously increase the firm's output and use of capital and hence causes the overestimation of the coefficient of capital in OLS

¹¹ The correlation between other-industry concentration and the interaction term is 0.92, and the collinearity might more or less lower the statistical significance of the coefficient of the interaction term. On the other hand, the industrial concentration in certain industries that are strongly linked with the target industry might have some impact on firm productivity. The investigation of this mechanism, however, is beyond this paper.

estimates. Alternatively, if capital stock is poorly measured, the attenuation bias could become stronger after using fixed effects and thereby further reduce the coefficient of capital. The coefficient of employment rises under fixed effects, but the extent of changes is relatively small.¹²

4.2 Robustness

We test the robustness of our findings to alternative choices in the research design. We address: (1) the temporal scope of the industrial environment; (2) the spatial dimension of agglomeration economies; (3) the industrial scope of localization and urbanization economies; (4) the measures of agglomeration economies and the dominance; (5) the dominance of SOEs and Non-SOEs; and (6) consistent estimates of standard errors. The results of those tests are displayed in Table 3.

We find that adding lagged regional industrial environment does not change our conclusions. In column (10) we investigate what the empirical results would be if we assume it is lagged, not contemporary, industrial environment that affects firm productivity. So column (10) uses those 1-year lagged variables such as $H_{j,t-1}$, $L_{j,t-1}$, $H_{j,t-1} * L_{j,t-1}$ and $U_{j,t-1}$. We hence use firms that stay in the data for at least 3 years and it reduces the sample size from 1.17 to 0.94 million. Similar to column (6) in Table 2, column (10) suggests that industrial concentration limits localization economies. Column (11) includes both contemporary and 1-year lagged variables of regional industrial environment, and column (12) adds the 2-year lagged variables in regressions. The results suggest that these lagged variables are all statistically insignificant, and the coefficients of original variables remain largely unchanged. This does not necessarily mean that industrial concentration and agglomeration in the past have no influences at all on firm productivity at present. Our interpretation is that, after controlling the variables at present, adding lagged variables introduces little new information. Indeed, *R* square rises very slightly, from 0.8502 in column (6) to 0.8677 in column (11) and 0.8759 in column (12). We also find a very high correlation between the current and lagged variables. The correlation is 0.9106 between current value and 1-year lagged value for own-industry concentration and 0.8492 between the current value and the 2-year lagged value. This intertemporal correlation is 0.9553 and 0.9238 for localization agglomeration, and 0.9952 and 0.9889 for urbanization agglomeration. In results not reported here, we find that adding lagged other-industry concentration does not affect our results either.¹³

We reveal that capturing agglomeration economies outside the city does not alter our conclusions. We measure the dominance and agglomeration only within city proper or county-level city because the literature suggests agglomeration

¹² These interpretations are based on the assumptions that the log-linear function form is a good approximation for the production function, and that the coefficients of employment and capital inputs do not vary across observations. While these assumptions are often made in the empirical literature, investigating the degree of the approximation is beyond this paper.

¹³ Whether it is lagged industrial concentration, contemporary industrial concentration, or both, that influence firm productivity is an important question, and we leave it for future research.

Table 3 Robustness of estimation results

	Time scope		Spatial scope	Industrial scope	Measure of agglomeration	Measure of dominance	Dominance of SOEs	Cluster
	(10)	(11)						
<i>Dependent variable: log (value added)</i>								
Dominance in own industry: Herfindahl index on sales		0.0554 (0.0324)	0.0640 (0.0363)	-0.0002 (0.0298)	0.0141 (0.0196)	-0.1133 (0.0197)**		0.0316 (0.0458) (0.0302) (0.0393)
Dominance in own industry: Leading firm share						0.0032 (0.0336)		
Dominance of SOEs in own industry: Leading firm share							0.0469 (0.0390)	
Dominance of Non-SOEs in own industry: Leading firm share							-0.0004 (0.0349)	
Localization: Log (own-industry employment in city)		0.0232 (0.0039)**	0.0231 (0.0043)**	0.0213 (0.0037)**	0.0157 (0.0025)**	0.0222 (0.0040)**	0.0229 (0.0041)**	0.0252 (0.0055)** (0.0040)** (0.0056)**
Localization: Log (no. of own-industry firms)						0.0187 (0.0047)**		

Table 3 (continued)

	Time scope		Spatial scope	Industrial scope	Measure of agglomeration	Measure of dominance	Dominance of SOEs	Cluster
	(10)	(11)						
Interaction between localization and own-industry dominance		-0.0318 (0.0041)**	-0.0319 (0.0046)**	-0.0336 (0.0040)**	-0.0236 (0.0025)**	-0.0425 (0.0081)**	-0.0244 (0.0042)**	-0.0362 (0.0061)** (0.0045)** (0.0070)**
Interaction between localization and own-industry dominance of SOEs								-0.0226 (0.0051)**
Interaction between localization and own-industry dominance of Non-SOEs								-0.0256 (0.0044)**
Log (own-industry employment in rest of prefecture)				-0.0017 (0.0025)				
1-year lagged dominance in own industry: Herfindahl index	0.0447 (0.0266)	0.0270 (0.0275)	0.0272 (0.0326)					
1-year lagged localization: Log (own-industry employment)	0.0123 (0.0031)**	0.0044 (0.0032)	0.0024 (0.0037)					

Table 3 (continued)

	Time scope		Spatial scope		Industrial scope	Measure of agglomeration	Measure of dominance	Dominance of SOEs	Cluster
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Interaction between localization and own-industry dominance, 1-year lagged	-0.0115 (0.0034)**	-0.0011 (0.0034)	0.0008 (0.0039)						
2-year lagged dominance in own-industry: Herfindahl index			0.0046 (0.0274)						
2-year lagged localization: Log (own-industry employment)			0.0041 (0.0032)						
Interaction between localization and own-industry dominance, 2-year lagged			-0.0030 (0.0033)						
Urbanization: Log (all other manufacturing employment)	0.0261 (0.0151)	0.0210 (0.0158)	0.0053 (0.0142)	0.0366 (0.0164)*	0.0326 (0.0153)*	0.0229 (0.0041)**	0.0302 (0.0286) (0.0207)	0.0302 (0.0286) (0.0473)	
Urbanization: Log (no. of other manufacturing firms)						0.0954 (0.0175)**			

Table 3 (continued)

	Time scope		Spatial scope	Industrial scope	Measure of agglomeration	Measure of dominance	Dominance of SOEs	Cluster	
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Log (all other manufacturing employment in rest of prefecture)				-0.0527 (0.0180)*					
1-year lagged urbanization: Log (all other manufacturing employment)	0.0291 (0.0138)*	0.0102 (0.0138)	0.0095 (0.0145)						
2-year lagged urbanization: Log (all other manufacturing employment)			-0.0072 (0.0129)						
Log (employment)	0.5000 (0.0064)**	0.4990 (0.0064)**	0.5077 (0.0073)**	0.5166 (0.0060)**	0.5124 (0.0057)**	0.5132 (0.0057)**	0.5123 (0.0057)**	0.5122 (0.0057)**	0.5123 (0.0115)** (0.0099)** (0.0225)**
Log (capital)	0.1950 (0.0031)**	0.1948 (0.0031)**	0.2080 (0.0035)**	0.1973 (0.0029)**	0.1994 (0.0029)**	0.1994 (0.0029)**	0.1993 (0.0029)**	0.1992 (0.0029)**	0.1993 (0.0053)** (0.0040)** (0.0069)**
Industry-year, province-year, and firm/location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 3 (continued)

	Time scope		Spatial scope	Industrial scope	Measure of agglomeration	Measure of dominance	Dominance of SOEs	Cluster
	(10)	(11)						
R^2	0.8677	0.8677	0.8759	0.8512	0.8502	0.8502	0.8502	0.8502
Obs.	9,44,361	9,44,361	7,58,625	10,08,203	11,71,808	11,71,775	11,71,774	11,71,808

Observations are individual firms

**Statistical significance at 1% level, and *significance at 5% level

Numbers in parentheses are standard errors clustered at city-year level, except in column (16)

externalities are highly localized (Rosenthal and Strange 2003; Puga 2010). Nevertheless, we still examine agglomeration economies outside the city proper. A rough way to do so is to look at the entire prefecture. In column (12), we add own-industry employment and other-industry employment in the rest of prefecture (outside the city proper but inside the prefecture). Neither of them contributes to firm productivity. So we believe that city proper or county-level city is an appropriate spatial unit in examining agglomeration economies as well as the effect of industrial dominance.

Our results are robust to a finer industrial classification. We use 3-digit industrial code to construct variables of own-industry concentration, localization agglomeration and urbanization agglomeration. Column (13) displays the estimated results. We find that the coefficients of key variables in Column (13) are qualitatively similar to those in column (9), but the magnitudes have various changes. The effect of own-industry concentration declines by 35% as indicated by a smaller coefficient of the interaction term, but it remains statistically significant. The effect of other-industry concentration, though not reported, is small given the very low statistical significance. The elasticity of localization economies reduces to 1% at the mean level, while the elasticity of urbanization economies increases slightly to 3.66%.

Our conclusions remain the same to an alternative measure of agglomeration. While both number of firms and employment are used in the literature to proxy agglomeration, they might represent different sources of agglomeration. Henderson (2003) suggests that the number of firms may better capture knowledge spillovers and finds that what matters is the number of firms but not employment. In this research, we find that using employment or number of firms generates very similar results. Column (14) suggests a negative effect of own-industry concentration on localization economies, as well as strong evidence for localization and urbanization economies. A difference is that the coefficient of own-industry concentration becomes negative and statistically significant. It implies that the concentration, besides retarding localization economies, could hurt firm productivity directly.

We examine if our estimates change with alternative indices for regional industrial dominance other than the Herfindahl index. In column (15), we measure the dominance by the largest firm's share of sales, the so-called leading firm share in the literature (Amato 1995). Estimated results suggest similar conclusions. The interaction term obtains a negative and significant coefficient, showing that a larger share of the own-industry leading firm leads to weaker localization economies. The coefficient of the own-industry leading firm share is small and statistically insignificant, implying that it does not directly affect firm productivity. The evidence for localization and urbanization economies are both present. In experiments not reported here, we find that the leading firm share in other manufacturing industries does not have any effects on firm productivity.

One may worry whether or not our results are driven by the presence of state-owned enterprises. Without a doubt, SOEs are larger, with an average employment of 770 compared to 223 of Non-SOEs and may affect agglomeration externalities via different mechanisms, saying local protectionism. So we include both the dominance of SOEs and the dominance of Non-SOEs in regressions by using the leading firm share. Estimated results in column (16) show that the two types of dominance exhibit quite similar effects. Neither of them directly affects firm productivity, as

indicated by the insignificance of their coefficients. Both types of dominance reduce localization economies, as indicated by the negative and significant coefficients of the interaction terms. The two coefficients of interaction terms, we note, are very close to each other. Therefore, we conclude that it is insignificant whether it is SOEs or Non-SOEs that dominate the local market.

Last but not least, we investigate whether different spatial levels at which standard errors are allowed to cluster change our conclusions. In our main results, all standard errors are clustered at the city-year level. Although we do not think that the errors are necessarily clustered at larger levels, we let standard errors cluster at city level, province-year level, and province level, respectively, and in this order we report their estimates of standard errors in parentheses below every coefficient in column (17). Compared to column (9), a notable difference is that the coefficient for urbanization agglomeration loses statistical significance. Localization agglomeration and the interaction term both have expected sign and remain highly significant. Our conclusion that local concentration reduces localization economies holds.

4.3 Results of subsamples

We test whether the effects vary across different firm sizes, different types of industries, different ownerships of firms, and different periods. Table 4 displays estimated results.

We examine whether relatively small firms are severely harmed by industrial dominance. We use firms smaller than 900 employees, approximately the bottom 95% of the total, to estimate coefficients in function (1).¹⁴ Estimated results in column (19)–(20) imply greater negative effects of the concentration for relatively small firms, since the coefficient of the interaction term between localization and own-industry dominance in column (20) is larger than that in column (6). Using some different cutoffs on employment generated similar findings.

We divide manufacturing firms into three types of industries, traditional light industries, traditional heavy industries and high-tech industries and find notable differences across sectors. We find that the concentration does not affect firm productivity in high-tech industries [columns (22) and (23)] but reduces localization economies in light and heavy industries [columns (18)–(21)]. The concentration has a stronger effect for light industry than for heavy industry, as implied by a larger coefficient (around -0.05) of the interaction term in the light industry.

We find the evidence for localization economies are not in high-tech industries but in light and heavy industries. This is indicated by a statistically insignificant coefficient of localization economies for high-tech industry (columns 22–23), and a positive and significant coefficient of localization economies for both light and heavy industries (columns 18–21). At the mean level, the elasticity of localization

¹⁴ We believe the industrial dominance is largely determined by the largest 5% firms. Among firms bigger than 900 employees, the average employment is 2762 and the median is 1509. For comparison, among firms no larger than 900 employees, the average employment is 164 and the median is 102.

Table 4 Estimated results for subsamples

	Small firms		Light industries		Heavy industries		High-tech Industries		State-owned firms		Non-state-owned firms		2000–2003		2004–2007	
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)
<i>Dependent variable: log (value added)</i>																
Dominance in own industry: Herfindahl index on sales	-0.0186 (0.0299)	-0.0161 (0.0304)	0.0651 (0.0478)	0.0662 (0.0482)	0.0615 (0.0413)	0.0655 (0.0417)	-0.0973 (0.1040)	-0.1089 (0.1034)	-0.0407 (0.0837)	-0.0377 (0.0837)	0.0110 (0.0309)	0.0130 (0.0314)	-0.0088 (0.0546)	-0.0079 (0.0549)	0.0319 (0.0535)	0.0390 (0.0541)
Localization: Log (employment in own industry)	0.0229 (0.0036)**	0.0238 (0.0037)**	0.0345 (0.0060)**	0.0355 (0.0061)**	0.0272 (0.0050)**	0.0283 (0.0050)**	0.0190 (0.0142)	0.0180 (0.0140)	0.0040 (0.0108)	0.0046 (0.0108)	0.0235 (0.0037)**	0.0245 (0.0038)**	0.0166 (0.0064)**	0.0168 (0.0065)**	0.0342 (0.0069)**	0.0360 (0.0070)**
Interaction between localization and own-industry dominance	-0.0380 (0.0039)**	-0.0384 (0.0040)**	-0.0513 (0.0061)**	-0.0519 (0.0061)**	-0.0376 (0.0055)**	-0.0382 (0.0054)**	-0.0239 (0.0135)	-0.0225 (0.0134)	-0.0186 (0.0107)	-0.0189 (0.0106)	-0.0368 (0.0041)**	-0.0370 (0.0041)**	-0.0348 (0.0067)**	-0.0348 (0.0067)**	-0.0466 (0.0073)**	-0.0478 (0.0074)**
Dominance in other manufacturing industries: Herfindahl index on sales	-0.1489 (0.3676)	0.4149 (0.4251)	-0.2252 (0.4152)	-1.2502 (0.8148)	0.3628 (0.4231)	-0.3010 (0.4312)	-0.1487 (0.5156)	0.1496 (0.6197)								

Table 4 (continued)

	Small firms		Light industries		Heavy industries		High-tech Industries		State-owned firms		Non-state-owned firms		2000–2003		2004–2007	
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)
Urbanization: Log (employment in other manufacturing industries)	0.0273 (0.0175)	0.0294 (0.0161)	0.0215 (0.0195)	0.0206 (0.0183)	0.0231 (0.0181)	0.0248 (0.0165)	0.0698 (0.0277)*	0.0805 (0.0248)**	-0.0263 (0.0232)	-0.0300 (0.0202)	0.0267 (0.0183)	0.0303 (0.0171)	-0.0138 (0.0272)	-0.0131 (0.0254)	0.0773 (0.0270)**	0.0740 (0.0243)**
Interaction between urbanization and other-industry dominance	-0.0145 (0.0386)		-0.0772 (0.0462)		-0.0008 (0.0421)		0.0964 (0.0805)		-0.0488 (0.0431)		-0.0029 (0.0447)		0.0024 (0.0531)		-0.0494 (0.0650)	
Log (employment)	0.5044 (0.0055)**	0.5043 (0.0055)**	0.4881 (0.0074)**	0.4880 (0.0074)**	0.5113 (0.0059)**	0.5112 (0.0059)**	0.5842 (0.0127)**	0.5843 (0.0127)**	0.5331 (0.0163)**	0.5331 (0.0163)**	0.5018 (0.0057)**	0.5017 (0.0057)**	0.4581 (0.0098)**	0.4581 (0.0098)**	0.4417 (0.0085)**	0.4418 (0.0085)**
Log (capital)	0.1935 (0.0029)**	0.1936 (0.0029)**	0.1955 (0.0038)**	0.1956 (0.0038)**	0.2018 (0.0034)**	0.2019 (0.0034)**	0.1844 (0.0070)**	0.1844 (0.0070)**	0.2337 (0.0112)**	0.2337 (0.0112)**	0.1942 (0.0030)**	0.1943 (0.0030)**	0.1864 (0.0051)**	0.1864 (0.0051)**	0.1851 (0.0042)**	0.1852 (0.0042)**
Industry-year, province-year, and firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.8229	0.8229	0.8523	0.8522	0.8426	0.8426	0.8983	0.8983	0.9054	0.9054	0.8373	0.8373	0.8791	0.8791	0.8830	0.8830
Obs.	11,12,202	11,12,220	3,97,270	3,97,378	6,76,503	6,76,511	91,855	91,855	1,24,343	1,24,354	10,40,066	10,40,073	4,12,883	4,12,895	7,25,596	7,25,600

Observations are individual firms

Numbers in parentheses are robust standard errors clustered at the city-year level

**Statistical significance at 1% level, and *significance at 5% level

economies to value added in the light industry is 2.58%. This is higher than that of heavy industries, which equals 2.15%.

We find the evidence for urbanization economies are in high-tech industries, but not in light industries or heavy industries. Columns (22) and (23) show positive and significant coefficients of urbanization economies in high-tech industries, where the elasticity of urbanization economies lies between 7 and 8%, significantly higher than the ones we get when using all manufacturing firms (columns 8–9). Columns (18)–(21) report positive but insignificant coefficients of urbanization economies in light and heavy industries. It implies that the evidence for urbanization economies we found when using all manufacturing firms is mainly driven by high-tech industries.

The effects of employment and capital inputs change moderately across the three types of industries. The light and heavy industries have quite similar elasticities of employment and capital inputs, which vary between 49–51 and 19–20%, respectively. These two sectors constitute 91.6% of firm samples. The production in high-tech industries relies more on employment and less on capital, as compared to two other sectors. Its elasticity of capital input, 18.4%, is very close to those of light and heavy industries, while the elasticity of employment is approximately 14% higher than heavy industries and 20% higher than light industries. The assumption of same input elasticities across sectors that we impose when pooling all manufacturing sectors together to estimate function (1) does not seem very unrealistic. For comparison, Henderson (2003) reports similar coefficients for employment and capital across high-tech and machinery sectors.

We also re-estimate the empirical model by using SOEs and Non-SOEs separately. SOEs in China have different objectives from Non-SOEs (e.g., social security and macroeconomic management), and are often criticized for low efficiency. According to results in columns (24) and (25), SOEs are almost independent to the regional industrial environment. Industrial concentration does not affect the productivity of SOEs, neither does localization nor urbanization agglomeration. Results for Non-SOEs are quite similar to our results in columns (8) and (9), except for a lower level of significance of the coefficient of urbanization agglomeration.

Finally, we investigate whether the effect of industrial concentration changes over time during China's fast development. We divide the study period 2000–2007 into two periods, 2000–2003 and 2004–2007. Columns (28)–(31) report estimated results for the two periods separately. We find that own-industry concentration reduces localization economies in both periods and its effect slightly increases in absolute value over time (the coefficient of the interaction term changes from -0.0348 to approximately -0.047). Other-industry concentration exhibits no effects in either periods, similar to what we have previously identified. In addition, columns (28)–(31) provide evidence that agglomeration economies increase over time. The elasticity of localization economies, at the mean level, increases from approximately 1% in 2000–2003 to approximately 2.7–2.9% in 2004–2007. Urbanization economies obtain no evidence in 2000–2003 but become quite strong in 2004–2007, with an elasticity of 7.4%. These findings imply that agglomeration externalities act as a growing source of improved firm productivity and economic development in China.

5 Conclusions and final remarks

This paper investigates whether a concentrated regional industrial structure/concentration affects firm productivity. Using China's firm-level data from 2000 to 2007, we find no negative impacts of regional industrial structure directly on a firm's output, but we find that regional industrial structure matters for localization economies. More specifically, cities in which industrial sectors are more dominated by a few large firms in own sector tend to have weaker localization economies. The magnitude of the effect of regional industrial structure on localization agglomeration is not trivial, as revealed by the estimated coefficient of the interactive term. For instance, a one standard deviation increase in the own-industry concentration will lead to a reduction of approximately 40% on the elasticity of localization economies. We conclude that regional industrial structure has little impact on urbanization economies.

Our conclusions are robust with different measures of regional industrial structure (the Herfindahl index and leading firm share), with different proxies on agglomeration, with different industrial, intertemporal and spatial scopes of agglomeration externalities, and with subsamples by ownership, by period, and by industrial sector. A break-down of industrial sectors reveals that the effect of regional industrial structure on localization agglomeration is subject to types of industries. Specifically, we found that (1) regional industrial concentration negatively affects light and heavy industries, but does not affect high-tech firms; and (2) high-tech firms exhibit urbanization economies but not localization economies.

Our findings suggest two important policy implications. First, as firms are less productive in cities where their own industry is dominated by a few large firms, China should promote a more competitive industrial structure featuring less dominance of large firms. This is particularly important because both light and heavy industries tend to be negatively affected by industrial dominance while simultaneously constituting the backbone of the national economy.

Second, SOEs are much larger than non-SOEs, in terms of both average size and average output. This implies that the marketization or privatization of SOEs could help to increase the competitiveness of the regional industrial structure (e.g., reducing the staff, spin-off, etc.), generating more positive externalities in addition to its own merits such as motivating management and reducing agency cost. Finally, small cities are more vulnerable to a concentrated industrial structure. This down-side effect of large firms on local/regional economy should be considered when seeking attraction of large investment in an industry.

References

- Acs ZJ (1992) Small business economics: a global perspective. *Challenge* 35(6):38–44
- Acs ZJ, Audretsch DB (1988) Innovation in large and small firms: an empirical analysis. *Am Econ Rev* 78:678–690
- Amato L (1995) The choice of structure measure in industrial economics. *Q J Bus Econ* 34:39–52
- Audretsch DB (2001) Research issues relating to structure, competition, and performance of small technology-based firms. *Small Bus Econ* 16(1):37–51

- Batisse C (2002) Dynamic externalities and local growth: a panel data analysis applied to Chinese provinces. *China Econ Rev* 13(2):231–251
- Booth DE (1986) Long waves and uneven regional growth. *South Econ J* 53(2):448–460
- Brandt L, Van Biesebroeck J, Zhang Y (2014) Challenges of working with the Chinese NBS firm-level data. *China Econ Rev* 30:339–352
- Carree MA, Thurik AR (1999) *Industrial structure and economic growth*. Cambridge University Press, Cambridge, pp 86–110
- Chen BR, Wu M (2014) Industrial agglomeration and employer compliance with social security contribution: evidence from China. *J Reg Sci* 54(4):586–605
- Chinitz B (1961) Contrasts in agglomeration: New York and Pittsburgh. *Am Econ Rev* 51(2):279–289
- Ding C (2007) Policy and praxis of land acquisition in China. *Land Use Policy* 24(1):1–13
- Ding C (2013) Transportation development, regional concentration and economic growth. *Urban Stud* 50(2):312–328
- Ding C, Li Z (2019) Size and urban growth of Chinese cities during the era of transformation toward a market economy. *Environ Plan B: Urban Anal City Sci* 46(1):27–46. <https://doi.org/10.1177/2399808317696072>
- Ding C, Niu Y (2019) City size, market fragmentation and firm productivity for manufacturing in China. *Reg Sci Urban Econ* 74:81–98. <https://doi.org/10.1016/j.regsciurbeco.2018.11.007>
- Drucker J (2013) Industrial structure and the sources of agglomeration economies: evidence from manufacturing plant production. *Growth Change* 44(1):54–91
- Drucker J, Feser E (2012) Regional industrial structure and agglomeration economies: an analysis of productivity in three manufacturing industries. *Reg Sci Urban Econ* 42(1):1–14
- Duranton G, Puga D (2001) Nursery cities: urban diversity, process innovation, and the life cycle of products. *Am Econ Rev* 91(5):1454–1477
- Enright MJ (1995) Organization and coordination in geographically concentrated industries. In: Lamoreaux NR, Raff DM (eds) *Coordination and information: historical perspectives on the organization of enterprise*. University of Chicago Press, Chicago, pp 103–146
- Feser EJ (2002) Tracing the sources of local external economies. *Urban Stud* 39(13):2485–2506
- Fu S, Hong J (2011) Testing urbanization economies in manufacturing industries: urban diversity or urban size? *J Reg Sci* 51(3):585–603
- Gai EQ, Zhu X, Cheng MW, Shi QH (2015) Factor market's distortion, markup and TFP. *Econ Res J* 50(05):61–75
- Glaeser EDL, Kerr WIR (2009) Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain? *J Econ Manag Strategy* 18(3):623–663
- Glaeser EL, Kallal HD, Scheinkman JA, Shleifer A (1992) Growth in cities. *J Polit Econ* 100(6):1126–1152
- Gopinath M, Pick D, Li Y (2004) An empirical analysis of productivity growth and industrial concentration in US manufacturing. *Appl Econ* 36(1):1–7
- Henderson JV (2003) Marshall's scale economies. *J Urban Econ* 53(1):1–28
- Henderson V, Kuncoro A, Turner M (1995) industrial development in cities. *J Polit Econ* 103(5):1067–1090
- Holz C (2008) How can a subset of industry produce more output than all of industry. Retrieved from <http://carstenholz.people.ust.hk/CarstenHolz-industry-stats-07-web-27Nov08.pdf>
- Hu C, Xu Z, Yashiro N (2015) Agglomeration and productivity in China: firm level evidence. *China Econ Rev* 33(2):50–66
- Jacobs J (1969) *The economy of cities*. Random House, New York
- Ke S (2010) Agglomeration, productivity, and spatial spillovers across Chinese cities. *Ann Reg Sci* 45(1):157–179
- Lee CY (2005) A new perspective on industry R&D and market structure. *J Ind Econ* 53(1):101–122
- Lee BS, Jang S, Hong SH (2010) Marshall's scale economies and Jacobs' externality in Korea: the role of age, size and the legal form of organisation of establishments. *Urban Stud* 47(14):3131–3156
- Levin RC, Cohen WM, Mowery DC (1985) R&D appropriability, opportunity, and market structure: new evidence on some Schumpeterian hypotheses. *Am Econ Rev* 75(2):20–24
- Lin HL, Li HY, Yang CH (2011) Agglomeration and productivity: firm-level evidence from China's textile industry. *China Econ Rev* 22(3):313–329
- Long C, Zhang X (2012) Patterns of China's industrialization: concentration, specialization, and clustering. *China Econ Rev* 23(3):593–612

- Lu J, Tao Z (2009) Trends and determinants of China's industrial agglomeration. *J Urban Econ* 65(2):167–180
- Moulton BR (1986) Random group effects and the precision of regression estimates. *J Econom* 32(3):385–397
- Porter ME (1998) Clusters and the new economics of competition. *Harvard Bus Rev* 76(6):77–90
- Puga D (2010) The magnitude and causes of agglomeration economies. *J Reg Sci* 50(1):203–219
- Rantisi NM (2002) The competitive foundations of localized learning and innovation: the case of women's garment production in New York City. *Econ Geogr* 78(4):441–462
- Rosenthal SS, Strange WC (2003) Geography, industrial organization, and agglomeration. *Rev Econ Stat* 85(2):377–393
- Saxenian A (1994) Regional advantage: culture and competition in Silicon Valley and Route, vol 128. Harvard University Press, Cambridge
- Scherer FM, Ross D (1990) Industrial market structure and economic performance. Houghton Mifflin Company, Boston
- Staiger DO, Stock JH (1997) Instrumental variables regression with weak instruments. *Econometrica* 65(3):557–586
- Tang H (2011) Zhongguo Xianzhuang: Guoyou Longduan Qiye Ji Guanliao He Zibenjia Yu Yishen [Current situation in China: state-owned monopoly enterprises gather bureaucrats and capitalists in one]. Retrieved from <http://business.sohu.com/20111008/n321494833.shtml>
- Yang CH, Lin HL, Li HY (2013) Influences of production and R&D agglomeration on productivity: evidence from Chinese electronics firms. *China Econ Rev* 27:162–178
- Young A (2000) The Razor's edge: distortions and incremental reform in the People's Republic of China. *Q J Econ* 115(4):1091–1135

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.