EERI
Economics and Econometrics Research Institute

Testing Urbanization Economies in Manufacturing Industries: Urban Diversity or Urban Size?

Shihe Fu and Junjie Hong

EERI Research Paper Series No 38/2010

ISSN: 2031-4892

EERI
Economics and Econometrics Research Institute
Avenue de Beaulieu
1160 Brussels
Belgium

Tel: +322 299 3523
Fax: +322 299 3523
www.eeri.eu

Testing Urbanization Economies in Manufacturing Industries: Urban Diversity or Urban Size?

Shihe Fu
Research Institute of Economics and Management
Southwestern University of Finance and Economics
55 Guanghua Cun Street
Chengdu, 610074, China
E-mail: fush@swufe.edu.cn
and
Cheung Kong Graduate School of Business
Beijing, 100738, China

Junjie Hong
School of International Trade and Economics
University of International Business and Economics
10 East Huixin Street, Chaoyang District
Beijing, 100029, China
E-mail: hongjunjie@alumni.nus.edu.sg

May 2010

Abstract. Whether urbanization economies stem from urban diversity or urban scale is not clear in the literature. This paper uses the 2004 China manufacturing census data and tests simultaneously the effects of urban size and industrial diversity on firm productivity, controlling for localization economies and human capital externalities. We find that productivity increases with city size—but at a diminishing rate, and the city size effect becomes negative for cities with population over two million. Firms also benefit from industrial diversity, and the strength of such benefit increases with city size but decreases with firm size. The characteristics of agglomeration economies in a transition economy are also discussed.

Keywords: Urbanization economies; Industrial diversity; Jacobs externalities; City size

JEL Classifications: L60; R12; R30

Acknowledgements: The authors are very grateful to Mark Partridge, two anonymous referees, Richard Arnott, Ying Fang, Yuming Fu, Judy Pereira, V. Brian Viard, seminar participants at Tsinghua University, Central University of Finance and Economics, Xiamen University, Cheung Kong Graduate School of Business, the 13th Asian Real Estate Society Annual Meeting and International Conference, and the 2008 Chinese Economists Society Annual Meetings, for very helpful comments. The authors gratefully acknowledge the financial support from the Project 211 (Phase III) of the Southwestern University of Finance and Economics and the Project 211 (Phase III) of the University of International Business and Economics.

1. INTRODUCTION

Traditionally, urban agglomeration economies are categorized into two types: localization economies and urbanization economies. Localization economies refer to the economies of scale external to a firm but internal to an industry, and urbanization economies refer to the economies of scale external to an industry but internal to a city (Hoover, 1937). The argument that localization economies (or Marshallian externalities, in the dynamic context) stem from labor market pooling, input sharing, and information spillovers among firms in the same industry in a city has been well accepted (Rosenthal and Strange, 2001). However, whether urbanization economies stem from urban size or urban industrial diversity has been less clear.

In fact, the concept of urbanization economies is defined vaguely in the literature, and has been evolving gradually. At first, urbanization economies are defined as scale economies external to any industry and resulting from the general level of city economy (Hoover, 1937) and measured by city size (population). Henderson (1986) also believes that urbanization economies are determined by a city's size only, and not by its industrial composition. Since the 1990s, the focus has shifted gradually onto urban diversity. Henderson, Kuncoro, and Turner (1995) define urbanization economies as the benefit a firm obtains from both the overall scale and diversity of a city. Glaeser et al. (1992) use lack of industrial diversity as a measure of urbanization economies, or Jacobs externalities, in the dynamic context. Henderson (2003) also uses the terms urbanization economies and Jacobs externalities interchangeably. In this paper, we do not intend to argue what the precise definition of urbanization economies should be. Instead, we refer to the effect of city size on firm productivity as the city size (scale) effect, and the effect of urban diversity as Jacobs externalities, and test these two effects simultaneously, conditioning on other important controls.

Empirical evidence for the existence and magnitude of urbanization economies—in terms of both the city size effect and Jacobs externalities—is also mixed. Many empirical studies from the 1970s and 1980s provide evidence that there exist returns to scale in city size. For example, Sveikauskas (1975) finds that doubling a city size is associated with about 6% higher labor productivity in an average industry. Segal (1976) finds that labor productivity in metropolitan areas with population above two million is 8% higher than in the remaining cities. Moomaw (1985) finds that the productivity effect of city size averages 7% among manufacturing industries. All these studies use industry level data, without distinguishing between localization economies and urbanization economies, and fail to separate localization economies from the city size effect, possibly biasing the estimated city size effect upward.

Other studies that also use city population or population density as the measure of urbanization economies distinguish localization economies from urbanization economies, but reach different conclusions regarding the city size effect. Carlino (1979) uses the total number of reported manufacturing units in a metropolitan area to measure urbanization economies, and metropolitan area population for the urbanization diseconomies, and finds that the effect of population size is negative in 18 of the 19 two-digit industries in his sample. Using city or metropolitan area population to measure urbanization economies, Henderson (1986) finds little evidence of urbanization economies in manufacturing industries in the U.S. and Brazil; but Sveikauskas, Gowdy, and Funk (1988) find that urbanization economies do exist in the food processing industry, and Nakamura (1985) finds small urbanization economies in manufacturing industries in Japan. Baldwin et al. (2007) use the 1999 Annual Survey of Manufactures data in Canada and find that the productivity of manufacturing plants benefits from urbanization economies (measured by metropolitan area population): the elasticity is about 0.077. However, when using a panel data approach, Baldwin, Brown, and Rigby (2008) find that the coefficient of city size tends to be negative. A recent study by Broersma and Oosterhaven (2009) detects that urban size (measured by job density) has a positive effect on regional labor productivity but a negative effect on the growth of regional labor productivity in Netherlands in the 1990s, suggesting that congestion effect may have been dominant during this period.[1]

[1] All these studies, except Sveikauskas, Gowdy, and Funk (1988), also find localization economies at the same time. For a comprehensive review of city size effect on productivity, see Melo, Graham, and Noland (2009).

Existing empirical results show a similar, mixed pattern when industry diversity is used to measure urbanization economies. Various types of industry diversity indexes have been proposed. Glaeser et al. (1992) construct a lack of diversity index, the ratio of the employment in the five largest industries—excluding the industry in question—to a city's total employment, to measure Jacobs externalities, and find that industrial diversity promotes employment growth in industries. Another approach is to construct a sub-industry diversity index. For example, the manufacturing industry consists of many sub-industries. For any sub-industry in question, a diversity index can be constructed based on the information from all other sub-industries within the manufacturing industry in a city. Henderson, Kuncoro, and Turner (1995) explore this measure and find that mature industries do not benefit from urban diversity, but new high-tech industries do, while both types of industries benefit from Marshallian externalities. They conclude that Jacobs externalities help attract new industries while Marshallian externalities help retain existing industries. Henderson (2003) also explores this measure but finds little evidence of Jacobs externalities in the U.S. manufacturing industries. Using a variable similar to the Herfindahl index, in terms of value added or total output, to measure industrial diversity in Chinese provinces, Batisse (2002) finds that industrial diversity has a positive effect on local industrial growth, but Gao (2004) finds no such evidence.[2] By examining the effects of export-industry demand shocks on central city and suburban employment growth, Hollar (2006) confirms that urbanization economies do exist and stem from industrial diversity. [3]Table 1 summarizes the effect of city size and urban diversity on the performance of firms or industries, based on various empirical studies.

[Insert Table 1 here]

The mixed empirical results reviewed above suggest that the relation between urban size, urban diversity, and productivity is complex. This paper contributes to the literature by testing the effect of city size and urban diversity simultaneously. Different from almost all existing

---

[2] A few recent studies extend the concept of urban diversity to cultural diversity. Florida (2002) finds that the openness and diversity of an urban milieu is positively associated with the concentration of human capital and high-tech industries. Ottaviano and Peri (2005) construct a linguistic diversity index to measure cultural diversity in cities and find that overall cultural diversity has a positive effect on wages and employment of U.S.-born workers during 1970-1990. Fu (2007) finds Jacobs externalities in labor markets.

studies on either city size or industrial diversity, we control for localization economies, human capital externalities, and other important firm characteristics. Furthermore, we test how firms of different sizes benefit from agglomeration economies in cities of different sizes, which is rarely explored in empirical studies.

We use the firm level data drawn from the 2004 China manufacturing census to perform the test. Compared with data from developed countries, this data set offers a few advantages in testing agglomeration economies. First, due to administrative regulation on migration into cities, many Chinese cities are considered too small—except for a few cities that are too large (Au and Henderson, 2006), while, in developed countries, free migration has resulted in city sizes that maintain a steady state. Theories predict that for a city with a too small (large) size, the net marginal benefit of adding more population to the city should be positive (negative). Our empirical results should confirm this prediction.

Second, China's transition from a planned economy to a market economy provides a reasonable way to deal with sorting bias in cross-section models. Specifically, firms may move to locations with high concentrations of economic activities, biasing the estimated agglomeration coefficient upward. As China's economy has become increasingly market-oriented, young firms become more mobile than old firms, when confronting a location choice problem. Since young firms are more likely to sort into cities that are conducive to profit, estimated agglomeration economies' effects on young firms are more likely to be biased upward. If old firms are found to enjoy agglomeration economies, this will confirm the existence of agglomeration economies.[4] Furthermore, if the estimated agglomeration economies for old firms and young firms are not statistically different, then, we can conclude that the sorting bias is not serious.

Last but not least, as the ownership structure of firms in China has changed from dominantly state-owned to a more diverse composition over the course of the transition, it becomes interesting to test whether firms of different ownership enjoy or contribute to agglomeration economies differently. Do state-owned enterprises (SOEs) benefit from

[3] López and Südekum (2009) find that in Chilean manufacturing sector the cross-industry agglomeration effects (measured by the total number of plants in other industries nearby) take place only in a plant's upstream industries.

[4] Di Addario and Potacchini (2008) use a cross-section sample to detect the labor market agglomeration economies in Italy. They argue that the effects of agglomeration on wages are causal because there is almost no migration (sorting) across the labor markets in their sample. Similarly, in our sample, old firms are considered immobile, so the identified agglomeration economies for the old firm sample can be considered causal.

agglomeration economies since SOEs are more self-sufficient and less sensitive to market changes? Do domestic Chinese firms benefit from agglomeration economies differently from foreign firms in China? These questions will be examined in this paper.

We find that, in general, productivity increases with city size, but at a diminishing rate. The returns to city scale become decreasing beyond a certain threshold of population level (say, over one million). Firms also benefit from industrial diversity, and the strength of such benefit increases with city size but decreases with firm size. In most of the cities, only small firms benefit from industrial diversity; medium firms and large firms benefit little from industrial diversity. We also find that firms in general benefit from localization economies, but not from human capital externalities—except for foreign firms, and state-owned firms benefit little from agglomeration economies.

The rest of the paper is organized as follows: Section 2 briefly reviews the theories on the relationship between firm size, city size, and urban industrial diversity; Section 3 describes the data set; Section 4 specifies the econometric models to be estimated and Section 5 discusses identification issues; Section 6 presents the empirical results and Section 7 concludes.

2. THEORETICAL ANALYSIS

The different results of the effect of city size and industrial diversity on firm productivity reviewed in the introduction suggest that the relationships among urban scale, industrial structure, and firm size are interrelated and complex.

First, theoretically, the effect of city size on firm productivity is indeterministic. On the one hand, there are economies of scale associated with city size. Large city sizes can reduce the average cost of urban public goods. Such an effect can be considered pure economies of scale. Abdel-Rahman (2000) constructs a theoretical model and demonstrates that firms have incentive to concentrate in a city that provides public infrastructures, resulting in agglomeration economies. Large city sizes also imply statistical economies in product markets and labor markets. For example, a firm's demand is less variable if the number of buyers is large and buyers' demand is uncorrelated. On the other hand, there are diseconomies of scale associated with city size. When city sizes become larger, problems of congestion, high real estate rents, and other disamenities will arise. Therefore, city size may capture both productivity advantages and disadvantages. Theories on the optimal city size indicate that, when a city has an optimal population size, the forces of agglomeration economies are offset by the forces of disagglomeration economies, resulting in locally constant returns to scale (Arnott, 1979, 2004). Since a real city size may be less than or larger than the optimal size, the net marginal effect of city size can be positive or negative.

Second, the productivity effect of urban diversity results from different channels and depends on the size of firms. Primarily, urban diversity refers to the industrial diversity within a city. Most likely, there is no one who emphasizes the importance of urban diversity to urban growth more than does Jane Jacobs. This is why the benefit from urban diversity is dubbed Jacobs externalities, even from a static context. Jacobs (1961) stresses the important effect of urban diversity on city safety and city growth. In Jacobs (1969), she argues further that the growth of a city is determined by that city's ability to constantly add new works to old ones, and that urban diversity is crucial to information exchange and innovation in cities. Other benefits from urban diversity can be labor market pooling or statistical economies (Quigley, 1998; Duranton and Puga, 2000). Industrial diversity may reduce frictional unemployment and stabilize employment (Simon, 1988). Overall, the benefits from urban industrial diversity may capture statistical economies, knowledge spillovers, and labor market pooling. The productivity effect of urban diversity also depends on firm size. Jacobs (1961) argues that small firms benefit more from urban diversity in large cities because small firms depend more on external industrial environments for multiple intermediate inputs, while large firms are relatively self-sufficient. In addition, a high degree of industrial diversity fosters innovations, which is favorable to young and small firms (Chinitz, 1961; Duranton and Puga, 2000).

Finally, both specialization and diversification of cities are associated with city sizes. A few theoretical models explain why some cities specialize while some other cities diversify. A city may specialize in one good if localization economies exist but cross-industry benefits do not exist (Henderson, 1974), or if the production of that good has internal scale economies (Abdel-Rahman and Fujita, 1993). A city may diversify if economies of scope exist in that city (Abdel-Rahman and Fujita, 1993), or if intermediate differentiated inputs can be shared by multiple industries (Abdel-Rahman, 1990). Although city sizes may be closely related to industrial structure, two cities of the same size may have very different degrees of industrial diversity and may show very different productivity effects. Although large cities tend to host more diverse industries, city size does not necessarily represent diversity. Furthermore, industrial diversity of a small urban area may also have a productivity effect. Empirical evidence from Rosenthal and Strange (2003) indicates that industrial diversity at the zip code level has a positive effect on the birth and employment of new establishments.

In summary, various theories show that city size, localization economies, Jacobs externalities, and firm size are closely interrelated. This paper aims to test from what types of agglomeration economies firms of different sizes can benefit, in cities of different sizes.

3. DATA

The data used in this paper are drawn from the first economic census of China, conducted by the Chinese government from 2004 to 2005, and covering the entire universe of establishments in China. We obtain the firm level data of manufacturing industries from the National Bureau of Statistics of China (NBSC). The data contain detailed information on all manufacturing firms (over 1.3 million) at the end of 2004, including the geographic location, year of entry, ownership, total assets, total employment, employment by education, etc.

The data on city population and employment in cities by one-digit industry are from the *China Urban Statistical Yearbook* 2005 (published by the China Statistics Press). The city used in this paper is defined as city proper, including both an inner city area and suburban areas but excluding independent suburban counties. Although some independent suburban counties, officially speaking, belong to the administrative scope of a city, they do not actually function like an urban area. Therefore, the *China Urban Statistical Yearbook* advises researchers to use "city proper" when studying urban related issues. Statistically, Chinese cities are classified into five categories, according to their population sizes: small cities, with population less than 0.2 million; medium cities, population between 0.2 and 0.5 million; large cities, population between 0.5 and 1 million; extra-large cities, population between 1 and 2 million; and super-large cities, with population above 2 million persons.

Industrial employment data are not available for most of the small cities. Therefore, we select only firms located in cities with population equal to or larger than 0.2 million. Our sample includes 115 medium cities, 76 large cities, 30 extra-large cities, and 20 super-large cities.

According to the NBSC, manufacturing firms are classified into three categories, based on the total number of employees, total revenue, or total assets. We adopt the total employee criterion. Small firms have less than 300 employees, medium firms 300~2000 employees, and large firms more than 2000 employees.

In the sample, only non-state-owned firms with sales above five million Yuan (the unit of Chinese currency) and state-owned firms are mandated to report the value of all the intermediate inputs, and only for these firms can the value added be computed. These firms form a subsample that we call the value-added sample. The original sample is called the total-output sample, since all firms report total output. The value-added sample accounts for about 23.8% of the firms in the total-output sample. There are three potential problems in using these two samples and we will discuss them in detail in Section 5.

4. MODEL SPECIFICATIONS

In line with most of the existing studies, we adopt the production function approach to test urbanization economies. Specifically, a firm's production function is specified as

(1) $$Y_{ijk} = f(\mathbf{X}_{ijk})g(\mathbf{L}_{jk})h(\mathbf{U}_k),$$

where Y_{ijk} is the value added of the *i*th firm in a two-digit manufacturing industry *j* located in city *k*, $\mathbf{X}_{ijk}$ is a vector of the firm's inputs, $\mathbf{L}_{jk}$ is a vector of characteristics of industry *j* located in city *k*, and $\mathbf{U}_k$ is a vector of characteristics of city *k*. *f* is assumed to be a neoclassical production function; *g* and *h* are functions measuring localization economies and urbanization economies, and are assumed to be Hicks neutral to *f*.

Since individual production function can vary across industries and locations, a flexible production function (translog production function) is preferred. We adopt the Cobb-Douglas production function form, with a set of other important control variables, and use the translog production function as a robustness check. Corresponding to equation (1), under some simplified assumptions, the benchmark econometric model can be specified as

(2) $$\ln Y_{ijk} = \alpha'\mathbf{X}_{ijk} + \beta'\mathbf{L}_{jk} + \gamma'\mathbf{U}_k + \varepsilon_{ijk},$$

where ε_{ijk} is a disturbance term, and α, β, and γ are coefficient vectors to be estimated. The dependent variable, $\ln Y_{ijk}$, is the natural logarithm of the value added (measured in 1,000 Yuan) of firm *i*, where value added equals the value of total output minus the value of total intermediate inputs.

The vector of inputs, $\mathbf{X}_{ijk}$, includes the following three variables:

$\ln(Collegeemp)$: the natural logarithm of a firm's total number of employees with a college degree or above at the end of year 2004, proxy for a firm's high-quality labor input;

$\ln(Noncollegeemp)$: the natural logarithm of a firm's total number of employees with less than a college degree at the end of year 2004, proxy for a firm's low-quality labor input;

$\ln(Asset)$: the natural logarithm of the monetary value (in 1,000 Yuan) of all the economic resources that a firm owns or controls, proxy for a firm's capital stock.

To control for observed firm heterogeneities, we add a set of the following variables to the $\mathbf{X}$ vector:

Age: a firm's age, equals 2004 minus the opening year;

Age square: the square of *Age*, proxy for the life cycle of a firm's products;

Female: the percentage of a firm's employees that are females.

In addition, we include a set of dummy variables to control for different types of registration, equity holding status, upper levels of administration, and organization levels.[5]

The vector of characteristics of a two-digit industry j in city k, $\mathbf{L}_{jk}$, consists of two variables:

Indavedu: the ratio of the number of employees that have a college degree or above in industry j in city k, excluding the college educated employees of the firm in question, to the total employees of industry j in city k, excluding the employees of the firm in question. The effect of this variable is referred to as the Marshallian type of human capital externalities in industry j in city k (Fu, 2007).

Specialization: the degree of specialization of industry j in city k. It equals the total employees in industry j in city k divided by the total employment in city k.[6] The effect of this variable is

[5] Registration type refers to the organization form of capital enrolled, including 23 types, such as state owned, collectively owned, proprietary, domestic joint-stock, and foreign. Equity holding refers to whether or not the state holds dominant equity shares. Upper level administration refers to which level of government supervises the firm, such as the central government, provincial government, and municipal government. Organization level means the rank of a firm in the political hierarchy of the state, province, city, and county.

[6] The available data of city employment are unit employment by one-digit industry, which exclude self-employed workers. Since self-employed workers are of small proportion and distributed across different industries, we believe that the ratio of unit employment in manufacturing industry in a city to the total unit employment in a city is very close to the actual manufacturing employment share in a city. Therefore, for each two-digit manufacturing industry in the census data, we compute its share in the total manufacturing employment, then, multiply this share by the ratio of manufacturing unit employment to the total unit employment in a city, to obtain the specialization index.

commonly referred to as localization economies or Marshallian externalities. We also try using the total employment in the same industry and the total number of firms in the same industry as a proxy for localization economies, respectively, and the patterns of the coefficients are very similar. But these two variables are highly correlated with city size (the correlation coefficients are 0.54 and 0.61, respectively), therefore, to avoid multicollinearity, we decide to use the *Specialization* index.

To better control for unobserved industry-specific characteristics, we also add two-digit industry fixed effects to the model.

The vector of characteristics of city *k*, $\mathbf{U}_k$, consists of two variables:

$\ln(Population)$: the natural logarithm of non-agricultural population at the end of year 2004 in city *k*, capturing the scale effect of city size;

Urban diversity: equals one minus the Herfindahl index, in terms of the employment in one-digit industries in city *k*, reflecting the degree of industrial diversity in that city. Specifically,

$$Urban\ diversity = 1 - \sum_{m=1}^{M} \left(\frac{E_{mk}}{\sum_{m=1}^{M} E_{mk}} \right)^2,$$

where E_{mk} is the number of employees in a one-digit industry *m* in city *k*, and *M* is the total number of one-digit industries in city *k*. There are 19 one-digit industries in total, including agriculture, manufacturing, mining, public utility, wholesale and retail trade, real estate, construction, etc., and the employment data are from the *China Urban Statistical Yearbook* 2005.[7] The value of *Urban diversity* is between zero and one, and a value closer to one implies that city industries are more diverse.

Since the diversity of manufacturing industries has also been used in the literature (Henderson, Kuncoro, and Turner, 1995; Henderson, 2003), we also construct a manufacturing diversity index that equals one minus the Herfindahl index, in terms of employment in the two-digit manufacturing industries in city *k*:

$$Manu\ diversity = 1 - \sum_{j=1}^{J} \left(\frac{E_{jk}}{\sum_{j=1}^{J} E_{jk}} \right)^2,$$

[7] The employment here also refers to unit employment. As explained in footnote 6, we believe that omitting self-employment does not generate serious bias, and that this index reflects well the actual degree of industrial

where E_{jk} is the number of employees in a two-digit manufacturing industry *j* in city *k* in the census data, and *J* is the total number of two-digit manufacturing industries in city *k*.

Some cities are capitals of provinces or are directly under the central government. Such cities may be favored politically or have attractive amenities. To better control for unobserved regional differences and some unobserved city characteristics, we add province fixed effects and a dummy variable indicating the status of being a capital city or being directly under the central government.

Before we complete the specification of the benchmark model, one point worth noting is that the relationship between industrial specialization and industrial diversity is nonlinear, or is not completely opposite (in our sample the correlation coefficient between *Specialization* and *Urban diversity* is about -0.32). A city can have multiple specializations while keeping a relatively high degree of industrial diversity at the same time.

5. IDENTIFICATION ISSUES

There are three potential identification issues in estimating equation (2) using the value-added sample. The first issue is sample truncation. Value added is considered the precise measure of a firm's final output because it avoids double counting the value of intermediate inputs, but using the value-added sample omits a large number of small, non-state-owned firms. Since the sample truncation uses mixed criteria: ownership and revenue, the standard truncated regression technique does not fit this case. Therefore, in this paper we present only the benchmark results by using the value-added sample, but we also estimate models using the logarithm of total output as the dependant variable based on the total-output sample, as robustness checks. Since the correlation between the logarithm of value added and the logarithm of total output is very high (about 0.85), the overall patterns of the results using these two different approaches are similar, as expected.

The second issue is aggregation bias. A firm may have multiple operating plants that spread into different cities, but we are unable to access plant information. We have to assume that all the employees of a firm are located in the same city. This might create some aggregation bias. However, we believe that the aggregation bias is not serious because multi-unit firms were

---

diversity in a city.

surveyed at the location where the majority of their business was conducted.[8] Also, in the value-added sample, approximately 95% of firms are single-unit firms, and they account for about 80% of the aggregate manufacturing sales, output, and employment. We also estimate the models separately for single-unit and multi-unit firms, for a robustness check.

The third issue is sorting bias. Some unobserved firm characteristics may correlate with industry or city attributes, biasing the estimates of β and γ. For example, competition in large cities is tougher; therefore, more productive firms may disproportionally sort into large cities. A popular way to deal with the sorting bias is to use a panel data approach to control for unobserved firm heterogeneities by including firm fixed effects. Since the 2004 China manufacturing census is the first and only census in China as of the writing of this paper, there is no way to construct a panel data.[9] As we have discussed in the introduction, young firms are more mobile and flexible in deciding where to locate, while old firms are more likely to be constrained since they were established under a planned economy; therefore, there should be more sorting across locations for young firms. We estimate the models for young firms and old firms separately and test if the coefficients of agglomeration variables are statistically different. If the coefficients are not statistically different, we can conclude that the sorting bias is not a serious problem.[10]

In addition, we recognize that excluding small cities might cause some bias in estimating the city size effect, but we believe that the potential bias does not affect much the overall pattern of our results, since only a very small proportion of manufacturing firms (employment) are located in small cities.

6. RESULTS

Overall Results

Table 2 presents the benchmark model results. All the standard errors are adjusted by city-industry clusters.[11] The coefficients of ln(*Asset*), ln(*Collegeemp*), and ln(*Noncollegeemp*) in all model specifications are plausible and relatively stable. Since most of the other

[8] A firm's headquarters also benefits from agglomeration economies or concentration of headquarters (Davis and Henderson, 2008).
[9] The second economic census was conducted in 2009 but the data are not yet available for research.
[10] For a brief review of other methods dealing with sorting bias, see Puga (2010).
[11] A city is coded by a three-digit number, and a two-digit manufacturing industry is coded by a two-digit number. We use 100*city code+industry code to form city-industry clusters.

coefficients of firm characteristics are of expected signs and significance, and are not of our particular interest, we focus on only the coefficients of four agglomeration variables. We first use *Manu diversity* to measure urban diversity—and the coefficient is negative and insignificant, indicating that manufacturing diversity does not enhance firm productivity. *Manu diversity* correlates moderately with ln(*Population*) and *Specialization* (correlation coefficients are 0.42 and -0.40, respectively); this might cause some collinearity problems (the coefficient of ln(*Population*) is 0.035, but insignificant). In addition, this variable reflects only the diversity of manufacturing industries, and does not capture the overall industrial diversity of a city. Therefore, we decide to replace this variable with the *Urban diversity* variable.

Column 1 confirms that *Urban diversity* is a much better measure: The coefficients of three out of four agglomeration variables are all significant at the 1% level. Even after controlling for human capital externalities and the city size effect, in general, firms still enjoy localization economies and Jacobs externalities. Doubling a city size is associated with about a 4.5% increase in total value added, which is comparable to the 3%~8% found in the literature (Rosenthal and Strange, 2004). The semi-elasticity of the *Specialization* index is about 0.91 and is significant at the 1% level, indicating that there exist significant localization economies in manufacturing industries. The semi-elasticity of the diversity index is about 0.35 and is also significant at the 1% level, providing evidence for the existence of Jacobs externalities in cities.

Columns 2 and 3 of Table 2 estimate the benchmark model for single-unit and multi-unit firms. The results for multi-unit firms are somewhat different: the coefficients of *Specialization* and *Urban diversity* are positive, but not significant, possibly because of the reduced sample size, or possibly because the headquarters of a multi-unit firm can provide intermediate services for its subsidiaries in a more efficient way than by outsourcing (Ono, 2003) and therefore benefit less from agglomeration economies. The results for single-unit firms, however, are remarkably similar to the results of the pooled data, suggesting that the aggregation bias is not a serious problem for pooled regression. Columns 4 and 5 present the results by non-high-tech and high-tech firms.[12] While the results of non-high-tech firms are similar to the results of the pooled data, the results of high-tech firms are a bit surprising: Although they benefit from human capital externalities, they do not benefit from either localization economies or urban

[12] The National Bureau of Statistics of China defines the classification of high-tech industries in China. It is a much broader classification than the high-tech industries selected in Henderson (2003). However, the results are

diversity, in contrast to existing studies (Henderson, 2003).[13]

To address the sorting bias issue, we estimate the model for young firms (firm age less than or equal to two years) and old firms (firm age greater than two years). Since young firms are more likely to sort into particular locations when confronting the location choice, and old firms are more constrained, then, if the sorting bias is serious, we would expect different coefficient estimates for the agglomeration variables for these two samples. Columns 6 and 7 present the results for young firms and old firms, respectively. Both young firms and old firms benefit from location economies, Jacobs externalities, and city bigness. A joined F test (F test statistic is 0.59) shows that the coefficients of the four agglomeration variables from the old-firm sample are not statistically different from those of young firms. We obtain similar results when using 5 or 10 years as the threshold to classify young and old firms. Therefore, we conclude that the sorting bias is not a major concern here.[14]

A few studies propose using a flexible production function, specifically, the translog production function (Nakamura, 1985; Henderson, 1986). We also try using this function. The patterns of the coefficients of agglomeration variables are relatively similar. Since using the translog production function does not generate new insights for this research, we decide to keep the simple Cobb-Douglas function form. We also try using the logarithm of total output as the dependent variable and expand the sample to 545,849 observations. The pattern of the estimate results based on the total-output sample is similar to that based on the value-added sample. This is mainly because the correlation between the logarithm of total output and the logarithm of value added is very high (0.85). Using value added per worker or total output per worker as the dependent variable also generates similar patterns.[15] In summary, the pooled data results show that, in general, manufacturing firms do benefit from localization economies, Jacobs externalities, and city bigness, but benefit little from human capital externalities.

---

similar when the sample is restricted to the similar high-tech industries defined in Henderson (2003).

[13] In column 5, *Specialization* and *Urban diversity* are highly correlated (correlation coefficient is 0.59). After dropping *Specialization*, the coefficients of the other three agglomeration variables are positive, but not significant. On average, high-tech firms are younger, have more employees, have a higher proportion of college graduates, and are more likely to be located in super-large cities, compared with other firms. However, testing the microfoundation of agglomeration economies in high-tech industries is beyond the scope of this paper.

[14] In their meta-analysis of 729 estimates taken from 34 studies of the productivity effect of urban agglomeration economies, Melo, Graham, and Noland (2009) conclude that endogeneity bias of agglomeration economies may not be a major concern.

[15] The results for robustness checks are not reported here but are available upon request.

[Insert Table 2 here]

Results by City Size

Urban theories predict that medium cities tend to be specialized, and large cities tend to be diverse (Abdel-Rahman and Fujita, 1990; Henderson, 1997). Empirically, theories imply that firms in medium cities enjoy more localization economies, while those in large cities enjoy more urbanization economies. We estimate the benchmark model by city size. The results are reported in Table 3. Columns 1-4 estimate the models for firms located in cities with populations greater than or equal to 0.2, 0.5, 1, and 2 million persons, respectively. A clear pattern emerges: As we restrict our sample gradually to larger cities, both localization economies and Jacobs externalities become stronger and become the strongest in super-large cities. The estimate results by city size category (columns 4-7) also confirm this pattern. This pattern seems inconsistent with the theory that city types and city sizes are determined by different degrees of scale economies associated with different production activities (Henderson, 1974). Our suggestive interpretation is that as cities become larger, the density of population, employment, and human capital tends to be higher too, generating more intensive and frequent social interactions and stronger knowledge spillover effects across firms.[16]

The city size effect also increases as we restrict our sample to larger cities, but in super-large cities (column 4 of Table 3) it becomes negative and significant at the 1% level, indicating that, probably, the diseconomies from city bigness may be dominant. We will provide further evidence in a moment. Human capital externalities are positive and significant only in super-large cities, possibly because super-large cities generate the strongest knowledge spillover effects due to the most intensive social interactions. In addition, the same level of education degree from different universities signals different human capital quality and the variable *Indavedu* fails to control for the quality of college degrees. Since super-large cities impose less regulation on immigrants with high-quality human capital, the same proportion of college graduates may generate more significant knowledge spillovers in super-large cities.

---

[16] Using urban land area data from the *China Urban Statistic Yearbook* 2005, we compute the population density, total employment density, and college graduates (in manufacturing industries) density. All the three density indexes increase monotonically with city size categories. The raw correlation coefficients of population density, employment density, and college graduates density with city population are 0.18, 0.22, and 0.29, respectively.

[Insert Table 3 here]

Results by Firm Size

Jacobs (1961) argues that small firms benefit more from urban diversity because small firms rely more on the external industrial environment. Rosenthal and Strange (2003) find that total employment at small establishments in the same industry, at the zip code level, has a larger effect on births and employment of new establishments than does total employment at medium or large establishments, possibly because small establishments are more open and innovative. These ideas suggest that firms of different sizes might benefit from, or contribute to, agglomeration economies in different ways. Therefore, we estimate the benchmark model by firm size. Table 4 presents the results: Small firms benefit from three types of agglomeration economies: localization economies, Jacobs externalities, and city bigness; medium firms benefit from localization economies and city bigness; large firms do not benefit from agglomeration economies. Using the total-output sample yields a similar pattern, except that large firms benefit marginally from localization economies and city bigness. Note that medium and large firms do not benefit from urban diversity, possibly because they are more self-sufficient, as Jacobs argues.[17]

The fact that small firms benefit more from localization economies than do large firms is worthy of more discussion. The relation between firm size and localization economies has not been fully understood yet in the literature. Kim (1995) and Holms and Stevens (2002) find positive correlation between industry concentration and plant size. Wheeler (2006) provides further evidence that localization economies are positively associated with the size of plants, but not the number of plants. However, small plants that specialize in intermediate goods tend to concentrate heavily, such as the dress industry in New York City (Lichtenberg, 1960). Using Italian manufacturing data, Lafourcade and Mion (2007) find that large plants are more concentrated (clustering within narrow urban areas), while small plants are less concentrated

[17] To make use of the full sample information, we interact the firm size dummies with the four agglomeration variables, using small size as the reference. The coefficient of *Specialization* for small firms is 0.9184 and is significant at the 1% level; the coefficients of firm size dummies for medium and large firms interacting with *Specialization* are -0.5040 and -0.4297, respectively. This confirms that small firms benefit strongly from localization economies, compared with medium and large firms. The same coefficients of firm size dummy variables interacting with *Urban diversity* are 0.4657, -0.5308, and -1.2452, respectively, and all are significant at the 1% level, suggesting that small firms benefit more strongly from Jacobs externalities than do medium and large firms.

but are more agglomerated (co-located within wider areas and spatially auto-correlated). How plant size is related to industrial concentration and localization economies warrants further investigation.

[Insert Table 4 here]

Results by Firm Size and City Size

To provide a complete picture of the effects of different types of agglomeration economies on firms of different sizes in cities of different sizes, for each firm size type, we estimate the benchmark model by city size. Table 5 reports the results. The results are remarkably consistent with the findings in the previous tables. The first panel shows that small firms benefit strongly from both localization economies and urban diversity in super-large cities. Small firms also benefit from city bigness in medium cities, but too large a size (say, over two million in population) means net diseconomies. The second panel shows that medium firms benefit from localization economies in extra-large cities, but benefit just marginally from urban diversity in extra- and super-large cities; medium firms also benefit from city bigness, but oversize (more than two million) implies diseconomies. These pieces of evidence indicate that the effects of urban size and urban diversity operate in different ways. The third panel shows that large firms benefit little from agglomeration economies. Using the total-output sample, and the logarithm of total output as the dependent variable, generates qualitatively similar results (results are not presented).[18] We should point out that Table 5's results are suggestive since estimating models by both city size and firm size reduces sample size dramatically and makes estimation less precise.

[Insert Table 5 here]

It is worth noting that, in all three panels of Table 5, the coefficients of ln(*Population*) are

[18] Since many studies have used industry level data or studied only a few particular two-digit manufacturing industries, we also estimate the benchmark model by two-digit industry. Most two-digit industries have just a few thousand observations, and the coefficients of agglomeration variables for most of the industries are positive, but not significant. When using the total-output sample and using the logarithm of total output as the dependent variable, the results show that the majority of industries enjoy the benefit from urban diversity and city bigness. There is also some evidence of localization economies. The results are in contrast to Nakamura (1985), where few urbanization economies are found in manufacturing industries in Japan.

negative for cities with population larger than two million, suggesting that super-large cities may have net diseconomies for all firms. To further investigate whether the scale effect is decreasing with city size, we add the square of ln(*Population*) to the benchmark model. Column 1 of Table 6 presents the result of the benchmark model using pooled data. The coefficient of the quadratic term of city size is positive, suggesting increasing returns to city scale in general. When we drop the medium cities, the coefficient of the quadratic term of city size is close to zero and not significant (column 2), suggesting locally constant returns to scale. Column 3 contains only extra-large and super large cities, the coefficient of the quadratic term of city size is negative and significant at the 1% level, suggesting diminishing returns to city scale; the same pattern holds for the super-large cities sample in column 4. Taken all together, Table 6 shows that in general the returns to city scale are diminishing as cities become larger.

[Insert Table 6 here]

7. AGGLOMERATION ECONOMIES IN A TRANSTION ECONOMY

China began its transition from a planned economy to a market-oriented economy in the early 1980s. During the transition process, firms face dramatic institutional reform and changes, which should have influenced both urban development and firms' location choices. One of the striking consequences of this transition is that the share of state-owned firms in the economy has been decreasing persistently. In this section, we test whether firms of different ownership or corporate governance structure benefit differently from agglomeration economies. Specifically, we estimate the benchmark model for subsamples of state-owned firms, non-state-owned firms, firms with or without a large state shareholder, domestic firms, and foreign firms.

Table 7 presents the results. Column 2 shows that state-owned firms benefit from neither localization economies nor urban diversity, but do benefit from city bigness. In contrast, non-state-owned firms benefit from three types of agglomeration economies.[19] One possible interpretation is that most of the SOEs were established under a planned economy, with the

[19] A joint F test (F=1.3) shows that the coefficients of column 2 are not statistically different from those of column 3, possibly because the small sample size of SOEs biases the estimates; however, using pooled data with the SOE dummy interacted with the four agglomeration economies variables (Chow test), a joint F test (F=11.8) shows that the coefficients of interaction terms are statistically different from zero, suggesting that SOEs and

responsibility of providing employees all kinds of welfare from cradle to grave, making SOEs more self-sufficient, heavily subsidized, and less competitive in markets. This might partially explain why state-owned firms do not benefit from localization economies and urban diversity. Columns 4 and 5 show that firms with a dominant state shareholder also gain no benefit from agglomeration economies, while firms with no dominant state shareholder benefit from three types of agglomeration economies. Column 6 shows that although in general domestic firms benefit from three types of agglomeration economies, they do not benefit from human capital externalities; however, foreign firms (column 7) benefit from human capital externalities, localization economies, and city bigness, but not from urban diversity, possibly because foreign firms are equipped with more advanced technology and management and therefore are less dependent on the diversity of the external environment.[20]

[Insert Table 7 here]

8. CONCLUSION

This paper uses the 2004 China manufacturing census data and tests the effects of urban industrial diversity and urban size on firm productivity, controlling for human capital externalities and localization economies. The results show that while firm productivity is positively associated with city size, the scale effect of city size decreases with city size. In general, firms benefit from urban diversity; small firms benefit more than do large firms, especially, small firms benefit strongly from urban diversity in super-large cities; however, large firms do not benefit from urban diversity, even in super-large cities. These findings are consistent with Jacobs's idea that small firms rely more on the diverse external environment.

We also find that larger cities tend to generate more significant localization economies and Jacobs externalities than do medium cities, possibly because larger cities generate stronger information spillover effects. State-owned firms and firms with a dominant state shareholder benefit much less from agglomeration economies, suggesting that state ownership may hinder a firm's ability to reap the benefits of external economies. We find little evidence of human

non-SOEs do benefit differently from agglomeration economies. The same discussion applies to the models in columns 4 and 5.

[20] The coefficients of columns 6 and 7 are jointly statistically different from each other.

capital externalities, except for foreign firms and for firms in super-large cities.

Our results can have policy implications for city governments, as they attempt to attract and retain firms. For example, to promote the growth of small firms, a local government can design policies to create a more diverse industrial environment. This is not a conflict to specialization. A city can be relatively diverse while hosting a few specialized industries. Likewise, although the existence of both localization economies and Jacobs externalities in Chinese cities can justify the boom of industrial parks or enterprise zones that many Chinese cities have seen during the past three decades, it also warns local governments that even an industry-specific external shock would have city-wide impact.

Our study can be easily extended by applying a panel data approach when such data become available. Our study does not identify empirically the exact mechanisms through which city size and urban diversity affect firm productivity. This topic would be an interesting avenue for future research.

REFERENCES

Abdel-Rahman, Hesham M. 1990. “Shared Inputs, Product Variety, and City Sizes,” *Journal of Regional Science*, 30, 359-374.

———. 2000. “Multi-firm City Versus Company Town: A Microfoundation Model of Localization Economies,” *Journal of Regional Science*, 40, 755-769.

Abdel-Rahman, Hesham M. and Masahisa Fujita. 1990. “Product Variety, Marshallian Externalities, and City Sizes,” *Journal of Regional Science*, 30, 165-183.

———. 1993. “Specialization and Diversification in a System of Cities,” *Journal of Urban Economics*, 33, 189-222.

Arnott, Richard. 1979. “Optimal City Size in a Spatial Economy,” *Journal of Urban Economics*, 6, 65-89.

———. 2004. “Does the Henry George Theorem Provide a Practical Guide to Optimal City Size?” *American Journal of Economics and Sociology*, 63, 1057-1090.

Au, Chun-Chung and J.Vernon Henderson. 2006. “Are Chinese Cities Too Small?” *Review of Economic Studies*, 73, 549-576.

Baldwin, John, Desmond Beckstead, Mark Brown and David Rigby. 2007. “Urban Economies and Productivity,” Economic Analysis Research Paper No. 045, Statistics Canada.

Baldwin, John, Mark Brown and David Rigby. 2008. “Agglomeration Economies: Microdata Panel Estimates from Canadian Manufacturing,” Economic Analysis Research Paper No. 049, Statistics Canada.

Batisse, Cecile. 2002. “Dynamic Externalities and Local Growth: A Panel Data Analysis Applied to Chinese Provinces,” *China Economic Review*, 13, 231-251.

Broersma, Lourens and Jan Oosterhaven. 2009. “Regional Labor Productivity in the Netherlands: Evidence of Agglomeration and Congestion Effects,” *Journal of Regional Science*, 49, 483 -51.

Carlino, Gerald. 1979. “Increasing Returns to Scale in Metropolitan Manufacturing,” *Journal of Regional Science*, 19, 363-373.

Chinitz, Benjamin. 1961. “Contrasts in Agglomeration: New York and Pittsburgh,” *American Economic Review*, 51, 279-289.

Davis, James and J.Vernon Henderson. 2008. “The Agglomeration of Headquarters,” *Journal of Urban Economics*, 38, 445-460.

Di Addario, Sabrina and Eleonora Patacchini. 2008. “Wages in the Cities: Evidence from Italy,” *Labor Economics*, 15, 1040-1061.

Duranton, Gilles and Diego Puga. 2000. "Diversity and Specialization in Cities: Why, Where and When Does It Matter?" *Urban Studies*, 37, 533-555.

Florida, Richard. 2002. "Bohemian and Economic Geography," *Journal of Economic Geography*, 2, 55-71.

Fu, Shihe. 2007. "Smart Café Cities: Testing Human Capital Externalities in the Boston Metropolitan Area," *Journal of Urban Economics*, 61, 86-111.

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

Glaeser, Edward, Hedi Kallal, Jose Scheinkman and Andrei Shleifer. 1992. "Growth in Cities," *Journal of Political Economy*, 100, 1126-1152.

Henderson, J.Vernon. 1974. "The Sizes and Types of Cities," *American Economic Review*, 64, 640-656.

———. 1986. "Efficiency of Resource Usage and City Sizes," *Journal of Urban Economics*, 19, 47-70.

———. 1997. "Medium Size Cities," *Regional Science and Urban Economics*, 27, 583-612.

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

Henderson, J.Vernon, Ari Kuncoro and Matthew Turner. 1995. "Industrial Development in Cities," *Journal of Political Economy*, 103, 1067-1090.

Hollar, Michael. 2006. "Industrial Diversity and Agglomeration Economies," HUD Working Paper No. 06-03.

Holmes, Thomas and John Stevens. 2002. “Geographic Concentration and Establishment Scale,” *Review of Economics and Statistics*, 84, 682-690.

Hoover, Edgar. 1937. *Location Theory and the Shoe and Leather Industries*. Cambridge, MA: Harvard University Press.

Jacobs, Jane. 1961. *The Death and Life of Great American Cities*. New York: Vintage Books.

———. 1969. *The Economy of Cities*. New York: Vintage Books.

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

Lafourcade, Miren and Giordano Mion. 2007. “Concentration, Agglomeration and the Size of Plants,” *Regional Science and Urban Economics*, 37, 46-68.

Lichtenberg, Robert. 1960. *One-Tenth of a Nation*. Cambridge, MA: Harvard University Press.

López, Ricardo and Jens Südekum. 2009. “Vertical Industry Relations, Spillovers, and Productivity: Evidence from Chilean Plants,” *Journal of Regional Science*, 49, 721-747.

Melo, Patricia, Daniel Graham and Robert Noland. 2009. “A Meta-Analysis of Estimates of Urban Agglomeration Economies,” *Regional Science and Urban Economics*, 39, 332-342.

Moomaw, Ronald. 1985. “Firm Location and City Size: Reduced Productivity Advantages as a Factor in the Decline of Manufacturing in Urban Areas,” *Journal of Urban Economics*, 17, 73-89.

Nakamura, Ryohei. 1985. “Agglomeration Economies in Urban Manufacturing Industries: A

Case of Japanese Cities," *Journal of Urban Economics*, 17, 108-124.

Ono, Yukako. 2003. "Outsourcing Business Services and the Role of Central Administrative Offices," *Journal of Urban Economics*, 53, 377-395.

Ottaviano, Gianmarco and Giovanni Peri. 2005. "Cities and Cultures," *Journal of Urban Economics*, 58, 304-337.

Puga, Diego. 2010. "The Magnitude and Causes of Agglomeration Economies," *Journal of Regional Science*, 50, 203 – 219.

Quigley, John. 1998. "Urban Diversity and Economic Growth," *Journal of Economic Perspective*, 12, 127-128.

Rosenthal, Stuart and William Strange. 2001. "The Determinants of Agglomeration," *Journal of Urban Economics*, 50, 191-229.

———. 2003. "Geography, Industrial Organization and Agglomeration," *Review of Economics and Statistics*, 85, 377-393.

———. 2004. "Evidence on the Nature and Source of Agglomeration Economies," in J.V. Henderson and J. Thisse (eds.), *Handbook of Regional and Urban Economics*, Vol.4. Elsevier North-Holland, pp. 2119-2171.

Segal, David. 1976. "Are There Returns to Scale in City Size?" *Review of Economics and Statistics*, 58, 339-350.

Simon, Curtis. 1988. "Frictional Unemployment and the Role of Industrial Diversity," *Quarterly Journal of Economics*, 103, 715-728.

Sveikauskas, Leo. 1975. "The Productivity of Cities," *Quarterly Journal of Economics*, 89, 393-413.

Sveikauskas, Leo, John Gowdy and Michael Funk. 1988. “Urban Productivity: City Size or Industry Size,” *Journal of Regional Science*, 28, 185-202.

Wheeler, Christopher. 2006. “Productivity and the Geographic Concentration of Industry: The Role of Plant Scale,” *Regional Science and Urban Economics*, 36, 313-330.

TABLE 1: The Effect of City Size and Urban Diversity on Productivity: A Summary

Paper	City size effect	Paper	Urban diversity effect
Sveikauskas (1975)	Positive	Glaeser et al.(1992)	Positive
Segal (1976)	Positive	Rosenthal and Strange (2003)	Positive
Moomaw (1985)	Positive	Batisse (2002)	Positive
Sveikauskas, Gowdy, and Funk (1988)	Positive	Henderson, Kuncoro, and Turner (1995)	Positive (new high-tech industry)
Nakamura (1985)	Positive		Negative (mature industries)
Baldwin et al.(2007)	Positive	Hollar (2006)	Positive
Henderson (1986)	Insignificant		
Carlino (1979)	Negative	Gao (2004)	Insignificant
Baldwin, Brown, and Rigby (2008)	Negative	Henderson (2003)	Little

TABLE 2: Benchmark Model Results

	1	2	3	4	5	6	7
	Pooled data	Single-unit	Multi-unit	Non-high-tech	High-tech	Young firms	Old firms
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
ln(Asset)	0.5247***	0.5235***	0.5279***	0.5121***	0.5784***	0.5024***	0.5249***
	100.01	97.97	31.85	96.28	44.52	58.63	93.35
ln(Collegeemp)	0.1433***	0.1399***	0.2016***	0.1323***	0.2456***	0.1000***	0.1551***
	31.55	30.39	13.16	30.09	15.31	13.92	32.32
ln(Noncollege-emp)	0.2385***	0.2372***	0.2728***	0.2574***	0.1518***	0.1888***	0.2528***
	40.08	39.79	14.92	40.70	13.30	22.27	39.88
Indavedu	0.0973	0.0946	0.0437	-0.1593	0.2325*	0.1203	0.0756
	0.97	0.92	0.18	-1.28	1.62	0.72	0.74
Specialization	0.9108***	0.9037***	0.5784	1.0927***	-0.3697	1.0069***	0.8796***
	3.34	3.23	1.51	3.44	-1.34	3.09	3.18
ln(Population)	0.0446***	0.0445***	0.0577*	0.0475***	0.0145	0.0496**	0.0434***
	3.17	3.07	1.83	3.17	0.47	2.33	3.13
Urban diversity	0.3475***	0.3431***	0.1681	0.3553***	-0.2776	0.4053***	0.3579***
	4.14	4.04	0.62	4.15	-1.20	3.08	4.26
Adjusted R^2	0.573	0.553	0.722	0.564	0.625	0.442	0.606
Sample size	129,947	123,365	6,582	117,813	12,134	30,198	99,749

Note. Dependent variable: ln(value added). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for provinces, capital cities and cities directly under the central government. Young firms are firms with age less than or equal to two years. Numbers below the coefficients are *t* statistics. Standard errors are adjusted by city-industry cluster. Superscripts "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels.

TABLE 3: Results by City Size

	1	2	3	4	5	6	7
	Firms in cities with population size (million)						
	≥0.2	≥0.5	≥1	≥2	1~2	0.5~1	0.2~0.5
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Indavedu	0.0973	0.1616	0.1987	0.4584***	-0.2588	0.2443*	0.1651
	0.97	1.43	1.39	2.91	-1.21	1.72	0.98
Specialization	0.9108***	1.0836***	1.298***	1.4837***	1.5260***	0.4497*	0.1867
	3.34	3.56	2.82	3.09	3.18	1.87	0.45
ln(Population)	0.0446***	0.0745***	0.1309***	-0.6850***	0.0147	-0.2342***	0.1778***
	3.17	4.19	3.66	-4.70	0.23	-3.80	2.92
Urban diversity	0.3475***	0.3334***	0.7349***	3.6160***	0.6430***	-0.3211***	0.0451
	4.14	3.44	3.95	6.29	3.93	-2.48	0.18
Adjusted R^2	0.573	0.571	0.581	0.577	0.599	0.546	0.606
Sample size	129,947	115,421	87,832	54,841	32,991	27,589	14,526

Note. Dependent variable: ln(value added). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for provinces, capital cities, and cities directly under the central government. Numbers below the coefficients are *t* statistics. Standard errors are adjusted by city-industry cluster. Superscripts "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels, respectively.

TABLE 4: Results by Firm Size

Variable	All firms Coefficient	Small firms Coefficient	Medium firms Coefficient	Large firms Coefficient
Indavedu	0.0973	0.0597	0.4123***	-0.6252
	0.97	0.56	2.66	-1.61
Specialization	0.9108***	0.8082***	0.7463***	-0.2895
	3.34	2.95	2.62	-0.52
ln(Population)	0.0446***	0.0396***	0.0837***	0.0385
	3.17	2.72	3.89	0.85
Urban diversity	0.3475***	0.4199***	0.0857	-0.5582*
	4.14	4.82	0.64	-1.66
Adjusted R^2	0.573	0.401	0.501	0.7 11
Sample size	129,947	107,320	20,681	1,946

Note. Dependent variable: ln(value added). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for provinces, capital cities, and cities directly under the central government. Numbers below the coefficients are *t* statistics. Standard errors are adjusted by city-industry cluster. Superscripts "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels, respectively.

TABLE 5: Results by Firm Size and City Size

	Firms in cities of population size (million)						
	≥0.2	≥0.5	≥1	≥2	1~2	0.5~1	0.2~0.5
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Small firms							
Indavedu	0.0597	0.1637	0.2627*	0.5456***	-0.3094	0.1708	0.0720
	0.56	1.37	1.74	3.25	-1.29	1.08	0.37
Specialization	0.8082***	1.0113***	1.3315***	1.5112***	1.4183***	0.2563	0.0138
	2.95	3.31	2.83	3.05	2.94	1.28	0.03
ln(Population)	0.0396***	0.0586***	0.1244***	-0.6649***	-0.0065	-0.2374***	0.1943***
	2.72	3.28	3.42	-4.52	-0.10	-3.62	2.86
Urban diversity	0.4199***	0.4358***	0.8057***	4.2550***	0.5986***	-0.2586*	0.1932
	4.82	4.43	4.22	7.32	3.55	-1.86	0.72
Adjusted R^2	0.401	0.403	0.425	0.427	0.434	0.339	0.408
Sample size	107,320	95,473	74,415	46,863	27,552	21,058	11,847
Medium firms							
Indavedu	0.4123***	0.3188*	0.0569	0.1503	0.0574	0.8766***	0.7738***
	2.66	1.82	0.28	0.60	0.18	3.07	2.64
Specialization	0.7463***	0.8801***	0.8295	0.7459	1.4527**	0.3563	-0.2288
	2.62	2.83	1.53	0.82	2.30	1.03	-0.40
ln(Population)	0.0837***	0.1469***	0.2139***	-0.6971***	0.1138	-0.2706**	0.1190
	3.89	5.45	3.74	-2.75	0.99	-2.30	1.19
Urban diversity	0.0857	-0.0063	0.3535	1.5607*	0.6080*	-0.5124***	-0.6091*
	0.64	-0.04	1.15	1.61	1.80	-2.74	-1.74
Adjusted R^2	0.501	0.504	0.513	0.513	0.525	0.470	0.507
Sample size	20,681	18,241	12,317	7,334	4,983	5,924	2,440
Large firms							
Indavedu	-0.6252	-0.7381	-0.1818	0.0126	0.1731	-1.4413*	-0.3366
	-1.61	-1.59	-0.30	0.02	0.20	-1.86	-0.42
Specialization	-0.2895	-0.3257	-0.9634	-3.4981**	0.6593	-0.5541	-1.6369*
	-0.52	-0.49	-0.81	-1.92	0.44	-0.69	-1.85
ln(Population)	0.0385	0.0613	0.1441	-0.1583	0.0172	-0.2693	0.0130
	0.85	1.14	1.04	-0.25	0.05	-1.19	0.05
Urban diversity	-0.5582*	-0.6215	-0.5581	0.8597	-1.6654	-1.0690**	-0.7195
	-1.66	-1.58	-0.65	0.35	-1.27	-1.94	-0.78
Adjusted R^2	0.711	0.709	0.725	0.749	0.716	0.667	0.783
Sample size	1,946	1,707	1,100	644	456	607	239

Note. Dependent variable: ln(value added). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for provinces, capital cities, and cities directly under the central government. Numbers below the coefficients are *t* statistics. Standard errors are adjusted by city-industry cluster. Superscripts "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels, respectively.

TABLE 6: Testing Diminishing Returns to City Scale

Variable	Firms in all cities Coefficient	Firms in cities above 0.5 million population Coefficient	Firms in cities above 1 million population Coefficient	Firms in cities over 2 million population Coefficient
Indavedu	0.1073	0.1619	0.1516	0.3054**
	1.07	1.43	1.09	2.02
Specialization	0.9129***	1.0839***	1.2214***	1.2519***
	3.36	3.57	2.55	2.45
ln(Population)	-0.4997*	0.0089	5.4636***	42.6761***
	-1.69	0.02	3.84	5.63
$[\ln(\text{Population})]^2$	0.0197*	0.0023	-0.1833***	-1.4391***
	1.83	0.11	-3.75	-5.72
Urban diversity	0.2949***	0.3311***	0.9960***	3.6078***
	3.22	0.49	4.82	6.96
Adjusted R^2	0.573	0.571	0.582	0.578
Sample size	129,947	115,421	87,832	54,841

Note. Dependent variable: ln(value added). Independent variables also include other firm characteristics, two-digit industry fixed effect, dummies for provinces, capital cities, and cities directly under the central government. Numbers below the coefficients are t statistics. Standard errors are adjusted by city-industry cluster. Superscripts "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels, respectively.

TABLE 7: Results by Ownership and Nationality

	1	2	3	4	5	6	7
Variable	All firms Coefficient	State-owned firms Coefficient	Non-state-owned firms Coefficient	State shareholder Coefficient	Non-state shareholder Coefficient	Domestic firms Coefficient	Foreign firms Coefficient
Indavedu	0.0973	-0.0537	0.1250	0.1401	0.1151	-0.0077	0.5494***
	0.97	-0.24	1.18	0.79	1.05	-0.07	3.73
Specialization	0.9108***	0.494	0.9492***	0.5294	0.9041***	0.8409***	1.0572***
	3.34	0.10	3.42	1.35	3.24	2.86	3.65
ln(Population)	0.0446***	0.0812**	0.0448***	0.0584*	0.0447***	0.0302**	0.0723**
	3.17	2.07	3.05	1.89	2.99	1.92	4.48
Urban diversity	0.3475***	0.4472	0.3132***	0.2262	0.3272***	0.3871***	0.1517
	4.14	1.47	3.65	0.96	3.74	3.80	1.46
Adjusted R^2	0.573	0.718	0.546	0.741	0.528	0.550	0.580
Sample size	129,947	8,582	121,365	12,496	117,451	94,131	35,816

Note. Dependent variable: ln(value added). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for provinces, capital cities, and cities directly under the central government. Numbers below the coefficients are t statistics. Standard errors are adjusted by city-industry cluster. Superscripts "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels, respectively.