Testing urbanization economies in manufacturing industries: urban diversity or urban size?

Fu, Shihe and Hong, Junjie

Southwestern University of Finance and Economics, University of International Business and Economics

August 2008

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

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

Abstract

Whether urbanization economies are caused by urban diversity or urban scale is not clear in regional and urban economics literature. Many empirical studies have used either city population size or urban industrial diversity to measure urbanization economies and have reached different conclusions. This paper argues that city size mainly captures the pure scale economies of urban public goods, and may generate net diseconomies when a city size is too large. Urban industrial diversity can also enhance firm productivity. Using the 2004 China manufacturing census data, we test simultaneously the effects of urban size and industrial diversity on firm productivity, controlling for localization economies and human capital externalities. We found that city size effect does exist, but too large a city size indicates net diseconomies. Firms also benefit from industrial diversity, and the strength of such benefit increases with city size but decreases with firm size. The overall results support Jacobs's idea that small firms benefit more from urban diversity.

Key words: Urbanization economies; Industrial diversity; Jacobs externalities; City size

JEL Classifications: L60; R12; R30

1 INTRODUCTION

Traditionally, urban agglomeration economies are categorized into two types: localization economies, referring to the economies of scale external to a firm but internal to an industry, and urbanization economies, referring to the economies of scale external to an industry but internal to a city (Hoover, 1937). Although Alfred Marshall's analysis that localization economies come from labor market pooling, input sharing, and information spillovers between firms in the same industry in a city has been widely accepted (Rosenthal and Strange, 2001), whether urbanization economies come from urban size or urban industrial diversity has been less clearer.

The concept of urbanization economies is, as a matter of fact, defined vaguely in the regional and urban economics 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, 1971) and measured by city size (population). Henderson (1986) also believed that urbanization economies are determined by only the size of the city, not its industrial composition. Since the 1990s the focus has been shifted gradually onto urban diversity. Henderson, Kuncoro, and Turner (1995) argued that urbanization economies mean the benefit a firm obtains from overall scale and diversity of a city. Glaeser et al. (1992) used "lack of industrial diversity" as a measure of urbanization economies, or Jacobs externalities, in the dynamic context, named after Jane Jacobs who emphasized the importance of urban industrial diversity on innovation and urban growth. Henderson (2003) also mixed urbanization economies with Jacobs externalities. In this paper, we do not intend to argue what the precise definition of urbanization economies should be. In stead, we call the effect of city population size on firm productivity city size (scale) effect, and the effect of urban diversity Jacobs externalities, and test simultaneously these two effects, conditioning on other controls.

Many empirical studies in the 1970s and 1980s have provided evidence that larger city size is associated with higher productivity. For example, Sveikauskas (1975) found that productivity is higher in larger cities: doubling a city size is associated with about 6% higher labor productivity in an average industry. Segal (1976) estimated an aggregate production function for 58 metropolitan areas and found that labor productivity in cities with population above two million is 8% higher, concluding that there exist returns to scale in city size or agglomeration effects.[1] Moomaw (1985) constructed a theoretical model incorporating city size choice in a firm's location choice problem. He found that the effect of city size on firm productivity averages 7% (the elasticity is 0.07) among manufacturing industries, and such external economies have been decreasing in larger cities during the 1970s. All these studies used industry level data, without the distinction between localization economies and urbanization economies, and failed to separate localization economies from city size effect.

Other studies that also used city population or population density as the measure of urbanization economies distinguished localization economies from urbanization economies, but reached different conclusions on the city size effect. Carlino (1979) used the total number of manufacturing reporting units in a metropolitan area to measure urbanization economies, and metropolitan area population the urbanization diseconomies, and found that the effect of population size is negative in 18 of the 19 two-digit industries in his sample. Henderson (1986) used industry-urban area level data in the USA and Brazil and measured urbanization economies by city population. He found little evidence of urbanization economies in manufacturing industries. Sveikauskas, Gowdy, and Funk (1988) measured urbanization economies in food processing industry by metropolitan area population, and found that such agglomeration effects do exist. Using the population in

[1] Both Sveikauskas and Segal's studies over-estimated the effect of city size due to improper model specifications or measurement errors in capital intensity (Moomaw, 1981).

densely inhabited districts as a measure, Nakamura (1985) found that small urbanization economies exist in manufacturing industries in Japan, and light industries tend to receive more urbanization economies than heavy industries do. The recent study by Baldwin et al. (2007) used metropolitan area population to measure urbanization economies and found that in Canada productivity of manufacturing plants benefits from urbanization economies: the elasticity is about 0.077.[2]

The different results of urban size effect suggest that the relationship between urban size and productivity is complex. 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) constructed a theoretical model and demonstrated 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 are large and buyers' demand are 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 the 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 the locally constant returns to scale (Arnott, 1979, 2004). Since real city sizes may not be optimal, the net effect of city size can be positive or negative. In addition, city sizes may not be closely related to industrial structure. Two cities of the same size may have very different degree of industrial diversity or specialization, thus, may show very different productivity effects.

[2] All these studies, except Sveikauskas, Gowdy, and Funk (1988), also found evidence on localization economies at the same time.

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 productivity effect. A piece of empirical evidence is from Rosenthal and Strange (2003) where industrial diversity at the zip code level has positive effect on the birth and employment of new establishments.

Urban diversity mainly refers to the industrial diversity within a city. Probably there is no one who emphasized the importance of urban diversity to urban growth more than Jane Jacobs did. This is why the benefit from urban diversity is loosely called Jacobs externalities, even from a static context. Jacobs (1961) stressed the important effect of urban diversity on city safety and city growth. Particularly, she argued that small firms benefit more from urban diversity in large cities because small firms depend more on external industrial environments while large firms are relatively self-sufficient. In her another book (1969) she argued further that the growth of a city is determined by that city's ability to constantly adding new works to old ones, and urban diversity is crucial in promoting urban innovation. Other benefit 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; Malizia and Ke, 1993). Overall, the benefits from urban industrial diversity may capture both statistical economies, innovation and knowledge spillovers, and labor market pooling.

A few studies constructed different indexes to test Jacobs externalities empirically, and the results are mixed. Glaeser et al. (1992) constructed a "lack of diversity" index, the ratio of the employment of the largest five industries excluding the industry in question to a city's total employment, to measure Jacobs externalities, and found that industrial diversity promotes employment growth in industries. Another approach is to construct a sub-industry diversity index. For example, manufacturing industry consists of many three-digit sub-industries. For any sub-industry in question, a diversity index can be constructed based on the information of all other sub-industries within the manufacturing industry in a city. Henderson, Kuncoro, and Turner (1995) explored this measure and found 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 concluded that Jacobs externalities help attract new industries while Marshallian externalities help retain existing industries. Henderson (2003) also explored this measure as well as the traditional city size measure in studying the urbanization-Jacobs economies in the US manufacturing and high-tech industries, and he found little evidence of Jacobs externalities.[3]

This paper contributes to the literature by using industrial diversity to measure urbanization-Jacobs economies while controlling for city size effect, localization economies, and human capital externalities. Different from Henderson (2003) and Rothensal and Strange (2003), we are looking at all manufacturing industries; different from Mommaw (1981), Glaeser et al. (1992), Henderson, Kuncoro, and Turner (1995), we are looking at the firm level; and most importantly, different from almost all existing studies on either city size or industrial diversity, we test the effect of city size and urban diversity simultaneously, conditioning on localization economies, human capital externalities, and other controls.

We use the 2004 China manufacturing census data to do the test. We find that in general, city size effect does exist, but too large a city size indicates net diseconomies. Firms also benefit from urban industrial diversity, and the

[3] Recent studies on urban diversity has extended the concept to cultural diviersity. For example, Florida (2002) found that the openness and diversity of urban milieu is positively associated with the concentration of human capital and high-tech industries. Ottaviano and Peri (2005) constructed a linguistic diversity index to measure cultural diversity in cities and found that overall cultural diversity has a positive effect on wages and employment of US-born workers during 1970-1990. Fu (2007) found Jacobs externalities in labor markets.

strength of such benefit increases with city size but decreases with firm size. Only small firms benefit from industrial diversity in all cities; medium firms and large firms benefit little from industrial diversity. The overall results support Jacobs's idea that small firms benefit more from urban diversity.

The next section describes the data. Section three specifies the econometric models to be estimated, and section four presents the results. The last section concludes.

2 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 obtained the firm level data of manufacturing industries from the State Statistical Bureau of China. The data set contains detailed information on all manufacturing firms (over 1.3 million) at the end of 2004, including the geographic location, year of entry, ownership, total asset, total employment, employment by education, etc.

The city used in this paper is defined as "city proper", not including the suburban counties. Statistically, Chinese cities are classified into five size categories according to their population sizes:

(1) Super-large cities: With a population above 2,000,000 persons;

(2) Extra-large cities: With a population between 1,000,000~2,000,000 persons;

(3) Large cities: With a population between 500,000~1,000,000 persons;

(4) Medium cities: With a population between 200,000~500,000 persons;

(5) Small cities: With a population less than 200,000 persons.

The data on city population and total employment by industry in cities are from the *China Urban Statistical Yearbook* 2005. The industrial employment data are not available for cities with population size less than 200,000.

Therefore, we select only firms located in cities with population equal to or larger than 200,000.

According to the State Statistical Bureau of China, manufacturing firms are classified into three categories, based on the total number of employees, total revenue, or total asset. 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.

One potential problem using this data set is that a firm may have multiple operating plants that spread in different cities. Since 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 a multi-unit firm was surveyed at the location where the majority of its business was conducted. Also in our sample approximately 97.34% of firms are single-unit firms. We also estimate the models separately for single-unit and multi-unit firms and the pattern of the results is very similar.

3 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

$$Y_{ijk} = f(X_{ijk})g(L_{jk})h(U_k), \tag{1}$$

where Y_{ijk} is the total output of the ith firm in a two-digit manufacturing industry j located in city k, X_{ijk} is a vector of the firm's inputs, L_{jk} is a vector of characteristics of industry j located in city k, and 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 control variables to control for unobserved firm, industry, and city characteristics, and use the translog production function as a robust check. Corresponding to (1), under some simplified assumptions, the benchmark econometric model can be specified as

$$\log Y_{ijk} = \alpha X_{ijk} + \beta L_{jk} + \gamma U_k + \varepsilon_{ijk}, \qquad (2)$$

where ε_{ijk} is a disturbance term, α, β, and γ are coefficient vectors to be estimated.

The vector of inputs, X_{ijk}, includes the following two variables:

Log(*Employee*): the logarithmic of a firm's total number of employees at the end of year 2004, proxy for a firm's labor input;

Log(*Asset*): the logarithmic of the monetary value of all the economic resources that a firm owns or controls, proxy for a firm's capital stock.

The key identification issue is that some unobserved firm characteristics may correlate with industry or city attributes, biasing the estimation of β and γ. To better control for firm heterogeneity, we add a set of the following variables to the 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, upper levels of administration, and organization levels.[4]

[4]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 the state has dominant equity shares or not. Upper level administration refers to which level of government supervises the firm, such as

The vector of characteristics of a two-digit industry j in city k, 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 to the total employees of industry j in city k, controlling for the quality of human capital stock in industry j in city k. The effect of this variable on output is also referred to as human capital externalities.[5]

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 commonly referred to as localization economies or Marshallian externalities.

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, U_k, consists of two variables:

Log(*population*): the logarithmic 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,

---

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.

[5]We also experiment with the percentage of a firm's employees that have a college degree or above to measure the average education level of a firm's labor force. This variable is not included in the model because it is higly correlated with *Indavedu* (the correlation is 0.41).

[6]The available data of employment by industry in a city is *unit* employment, which excludes 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 industries 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.

$$Urban\ diversity = 1 - \sum\nolimits_{m=1}^{M} \left(\frac{E_{mk}}{\sum\nolimits_{m=1}^{M} E_{mk}} \right)^2, \tag{3}$$

where E_{mk} is the number of employees in a one-digit industry m in city k, 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, real estate, construction, etc., and the employment data are from *China Urban Statistical Yearbook 2005.* [7] The value of $Urban\ diversity$ is between zero and one. As the value becomes closer to one, the city industries become more diverse.

Since the diversity of manufacturing industries has also been used in the literature (Henderson, Kuncoro, and Turner, 1995), we also construct a manufacturing diversity index to test its effect:

$Manu\ diversity$: equals one minus the Herfindahl index in terms of two-digit manufacturing industries in city k:

$$Manu\ diversity = 1 - \sum\nolimits_{j=1}^{J} \left(\frac{E_{jk}}{\sum\nolimits_{j=1}^{J} E_{jk}} \right)^2, \tag{4}$$

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.

To better control for unobserved regional differences and unobserved city characteristics, we add province fixed effects.[8] Some cities are capitals of provinces or directly under the central government. Such cities may be favored politically and we also create a dummy variable to control for such influences.

Before we complete the specification of the benchmark model, one point worth noting is that the relationship between industrial specialization and

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

[8] A city does not cross province boundaries.

industrial diversity is not linear or completely opposite. A city can have multiple specializations and can show specializations in a few industries and have a relatively high degree of diversity at the same time.

4 RESULTS

4.1 Overall results

Table 1 presents the benchmark model results. All the standard errors are adjusted by city-industry clusters. Since the coefficients of firm characteristics are of expected signs and significance, and are not of our particular interest, we report only the coefficients of agglomeration variables. Column 1 uses manufacturing diversity to measure urban diversity and the coefficient is negative and significant, indicating that manufacturing specialization enhances firm productivity. The variable $Manu\ diversity$ correlates moderately with *Log*(*population*) and *Specialization* (correlation coefficients are 0.4 and 0.28 respectively), this might cause come collinearity problem. In addition, this variable reflects only the diversity of manufacturing industries, not really the industrial diversity of a city. Therefore, we decide to replace this variable by $Urban\ diversity$ variable. Column 2 confirms that $Urban\ diversity$ is a much better measure, as the significance of all the agglomeration variables improves a lot. Even after controlling for human capital externalities and city size effect, in general, firms still enjoy localization economies and Jacobs externalities. Doubling a city size is associated with about 11% increase in total output, which is comparable to 8% in Segal (1976). The semi-elasticity of *Specialization* index is about 1.06, indicating there exist significant localization economies in manufacturing industries; while the semi-elasticity of diversity index is about 0.98, also indicating significant Jacobs externalities in cities.

Columns 3 and 4 in Table 1 estimate the benchmark model for single-unit and multi-unit firms. Although the results for multi-unit firms are somewhat different as expected, the results for single-unit firms are remarkably similar to the results of pooled data, suggesting that the "multi-unit bias" is not serious. The last two columns present the results by non-high-tech and high-tech firms.[9] While the results of non-high-tech firm is similar to results of pooled data, the results from high-tech firms are surprising: They do not benefit from any types of agglomeration economies, contrasting to existing studies.[10]

A few studies proposed to use flexible production function, specifically, translog production function (Nakamura, 1985; Henderson, 1986). We also try translog production function. All the results of agglomeration variables are pretty much similar (results are not reported here). Since using 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 output per worker as the dependent variable, and the pattern of the estimate results is very similar. In summary, the pooled data results show that in general, manufacturing firms benefit from human capital externalities, localization economies, city bigness, and Jacobs externalities.

[9] The classification of high-tech industries is defined by the State Statistical Bureau of China, which is much broader than high-tech industries selected in Henderson (2003).

[10] In column 6 variables *Specialization* and *Urban diversity* are highly correlated (correlation coefficient is 0.59). After dropping *Specialization*, the coefficients of other three agglomeration variables are positive but not significant. On average, high-tech firms are younger, have more employees, have higher proportion of college graduates, and are more likely to be located in super-large cities, compared with other firms. However, testing agglomeration economies in high-tech industries is beyond the scope of this paper.

TABLE 1 Benchmark model results

	1	2	3	4	5	6
	Pooled data	Pooled data	Single-unit	Multi-unit	Non-high-tech	High-tech
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Indavedu	0.1496	0.2131*	0.2070*	0.4078**	0.0680	0.2274
	1.29	1.82	1.76	2.03	0.54	1.30
Specialization	0.4497*	1.0642***	1.0598***	1.0317***	1.3903***	-1.4225***
	1.68	3.93	3.88	3.20	5.83	-3.24
Log(population)	0.0968***	0.1103***	0.1119***	0.0231	0.1124***	0.0621
	5.09	5.80	5.84	0.94	5.77	1.15
Urban diversity		0.9813***	0.9819***	0.5387***	1.0099***	-0.0909
		9.46	9.40	2.76	9.67	-0.29
Manu diversity	-0.4631***					
	-3.05					
Adjusted R^2	0.71	0.71	0.70	0.83	0.71	0.74
Sample size	545,849	545,849	531,720	14,129	510,025	35,824

Dependent variable: log (total output). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for capital cities and cities directly under the central government. Numbers below the coefficients are t statistics. Standard errors are clustered by city-industry combination. "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels.

4.2 Results by city size

Urban theories predict that medium cities tend to be specialized, and large cities tend to be diverse. Empirically, theories imply that firms in medium cities enjoy more localization economies, while in large cities enjoy more urbanization economies. We estimate the benchmark model by city size. The results are reported in Table 2. Columns 1-4 estimate the models for firms located in cities with population greater than or equal to 0.2, 0.5, 1, and 2 million, respectively. A clear pattern shows up: As we restrict our sample gradually to larger cities, localization economies attenuate while Jacobs externalities become stronger. Estimate results by city size type (columns 4-7) also confirm this pattern. Specifically, localization economies are the strongest in medium cities, while Jacobs externalities are the strongest in super-large cities. The city size effect in extra-large cities (column 5) is 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.

TABLE 2 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	0.2131*	0.2674**	0.1273	0.3165*	0.1258	0.4702***	-0.1496
	1.82	1.94	0.76	1.83	0.75	3.49	-1.20
Specialization	1.0642***	0.9565***	0.2417	0.6014	0.3360	1.1617***	1.9218***
	3.93	3.00	0.52	1.15	0.99	3.98	4.71
Log(population)	0.1103***	0.1525***	0.2626***	0.0682	-0.1739***	-0.0837	0.2136***
	5.80	5.89	5.91	0.59	-3.05	-1.57	4.35
Urban diversity	0.9813***	0.9491***	1.5136***	2.5013***	1.0839***	-0.2471	0.7312***
	9.46	7.61	7.80	5.47	6.80	-1.26	4.54
Adjusted R^2	0.71	0.71	0.72	0.72	0.73	0.70	0.73
Sample size	545,849	481,555	365,458	231,282	134,176	116,097	64,294

Dependent variable: log (total output). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for capital cities and cities directly under the central government. Numbers below the coefficients are t statistics. Standard errors are clustered by city-industry combination. "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels.

4.3 Results by firm size

Jacobs (1961) argued that small firms benefit more from urban diversity because small firms rely more on the external industrial environment. Rosenthal and Strange (2003) found that total employment at small establishments in the same industry at the zip code level has larger effect on births and employment of new establishments than does total employment at medium or large establishments, possibly because small establishments were 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 3 presents the results: Small firms benefit from all types of agglomeration economies; medium firms benefit from localization economies and city bigness; large firms also benefit from localization economies and city bigness, but the magnitudes attenuate by half compared with medium firms. Note that medium and large firms do not benefit from urban diversity, possibly because they are more self-sufficient, as Jacobs argued.[11]

That firms of different sizes all benefit from localization economies is worth more explanation. The relation between firm size and localization economies has not been fully understood yet in the literature. Kim (1995) and Holms and Stevens (2002) found positive correlation between industry concentration and plant size. Wheeler (2006) provided further evidence that

[11] To make use of the full sample information, we interact the firm size dummies with *Specialization* and *Urban diversity*, using small size as the reference. The coefficient of *Specialization* for small firms is 1.0161 and is significant at the 1% level; the coefficients of firm size dummies interacting with *Specialization* for medium and large firms are 0.5553 and 0.6321, respectively, but neither is significant. This confirms that small firms benefit strongly from localization economies. The coefficients of firm size dummy variable interacting with *Urban diversity* for small, medium, and large firms are 0.9891, -0.1806, and -0.2742, respectively, and all are significant at the 1% level, suggesting that small firms benefit more strongly from Jacobs externalities than medium and large firms do. In summary, when using firm size dummy variable interacting with agglomeration variables, the pattern of urban diversity effect is remarkably consistent with the results in Table 3.

localization economies are positively associated with the size of plants, but not the count of plants. However, small plants that specialize in intermediate goods tend to concentrate heavily, such as dress industry in New York (Lichtenberg, 1960). Using Italian manufacturing data, Lafourcade and Mion (2007) found that large plants are more concentrated (clustering within narrow urban areas), small plants are less concentrated 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.

TABLE 3 Results by firm size

Variable	All firms Coefficient	Small firms Cofficient	Medium firms Coefficient	Large firms Coefficient
Indavedu	0.2131*	0.2122*	0.1732	0.0193
	1.82	1.77	1.13	0.07
Specialization	1.0642***	1.0420***	1.2759***	0.6358*
	3.93	3.44	4.95	1.84
Log(population)	0.1103***	0.1107***	0.1119***	0.0639*
	5.80	5.71	5.03	1.71
Urban diversity	0.9813***	1.0404***	-0.0144	-0.4355
	9.46	9.89	-0.10	-1.40
Adjusted R^2	0.71	0.65	0.67	0.81
Sample size	545,849	520,918	22,917	2,014

Dependent variable: log (total output). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for capital cities and cities directly under the central government. Numbers below the coefficients are t statistics. Standard errors are clustered by city-industry combination. "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels.

4.4 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 4 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 localization economies in medium and large cities, while benefit strongly from urban diversity in super-large cities; small firms also benefit from city bigness in medium cities, but too large a size (say, over one million) means net diseconomies. The second panel shows that medium firms benefit from localization economies in large cities and above, but benefit from urban diversity only in super large cities; medium firms benefit from city bigness only in extra-large cities, and oversize (more than two million) implies net 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, except localization economies in medium cities.

TABLE 4 Results by firm size and city size

	Firms in cities of population size (Million)						
Small firms	≥ 0.2	≥0.5	≥1	≥ 2	1~2	0.5~1	0.2~0.5
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Indavedu	0.2122*	0.2689*	0.1436	0.3404**	0.1307	0.4494***	-0.1819
	1.77	1.90	0.83	1.95	0.74	3.27	-1.41
Specialization	1.0420***	0.8959***	0.1434	0.6421	0.1837	1.1594***	2.1092***
	3.44	2.50	0.27	1.23	0.50	3.95	5.12
Log(population)	0.1107***	0.1520***	0.2714***	0.0842	-0.1895***	-0.0811	0.2218***
	5.71	5.78	5.83	0.72	-3.24	-1.21	4.40
Urban diversity	1.0404***	1.0166***	1.5403***	2.4113***	1.0951***	-0.1956	0.7811***
	9.89	8.11	7.73	5.22	6.74	-1.22	4.73
Adjusted R^2	0.65	0.65	0.66	0.66	0.67	0.62	0.66
Sample size	520,918	459,526	351,154	222,811	128,343	108,372	61,392
Medium firms							
Indavedu	0.1732	0.1576	-0.1603	-0.1191	-0.2981	0.5556**	0.0796
	1.13	0.94	-0.81	-0.48	-0.92	2.09	0.23
Specialization	1.2759***	1.4451***	1.7192***	2.5451***	1.8479***	0.7088**	-0.1715
	4.95	5.33	4.45	2.88	4.04	2.01	-0.38
Log(population)	0.1119***	0.1612***	0.1314***	-0.7723***	0.1696*	-0.1915*	0.0206
	5.03	5.64	2.75	-3.58	1.86	-1.73	0.22
Urban diversity	-0.0144	-0.1180	0.4003	2.0931***	0.4625	-0.6734***	-0.4908
	-0.10	-0.70	1.39	2.46	1.48	-3.22	-1.62
Adjusted R^2	0.67	0.68	0.66	0.66	0.66	0.68	0.63
Sample size	22,917	20,262	13,170	7,809	5,361	7,092	2,655
Large firms							
Indavedu	0.0193	0.0111	0.2479	0.4280	0.3537	-0.8039	0.1768
	0.07	0.03	0.58	0.63	0.68	-1.28	0.27
Specialization	0.6358*	0.7563*	-0.4969	-1.1554	0.4154	1.6676***	0.1567
	1.84	1.71	-0.61	-0.68	0.43	3.05	0.24
Log(population)	0.0639*	0.0827*	-0.0582	-0.1016	0.0044	0.2255	0.2342
	1.71	1.88	-0.51	-0.18	0.02	0.82	1.14
Urban diversity	-0.4355	-0.4851	0.4112	1.3853	0.0440	-0.96*	-0.1951
	-1.40	-1.32	0.65	0.64	0.05	-1.79	-0.24
Adjusted R^2	0.81	0.81	0.82	0.82	0.83	0.78	0.83
Sample size	2,014	1,767	1,134	662	472	633	247

Dependent variable: log (total output). Independent variables also include other firm characteristics, two-digit industry fixed effects, dummies for capital cities and cities directly under the central government. Numbers below the coefficients are t statistics. Standard errors are clustered by city-industry combination. "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels.

It is worth noting that the coefficients of $Log(population)$ are not significant or even negative for cities with population larger than two million, suggesting that super-large cities may have no benefit or even net diseconomies to all firms. To further investigate this effect, we also add the square of $Log(population)$ to the benchmark model and estimate it by firm size in super-large cities. Table 5 presents the results and imputes the optimal city size. The imputed optimal city sizes range from 1,686,557 to 2,257,571 persons. However, the mean population of the 20 super-large cities is 4,006,440, the minimum is 2,044,600, the standard deviation is 2,204,273, and the maximum is 1,080,000 (Shanghai). This suggests that super-large cities are too large for manufacturing firms.[12]

[12] Au and Henderson (2006) found that many Chinese cities are undersized, but a few cities are significantly oversized.

TABLE 5 Testing optimal city size for firms in super-large cities

Variable	All firms Coefficient	Small firms Coefficient	Medium firms Coefficient	Large firms Coefficient
Indavedu	-0.0127	-0.0015	-0.2114	0.3625
	-0.06	-0.01	-0.87	0.51
Specialization	0.8224	0.8676	2.4764***	-1.2913
	1.51	1.58	2.77	-0.76
Log(population)	14.0126	15.2506*	16.1419	10.8758
	1.55	1.69	1.42	0.41
$[\text{Log(population)}]^2$	-0.4815	-0.5220*	-0.5629	-0.3717
	-1.61	-1.74	-1.49	-0.42
Urban diversity	0.7151**	0.6321*	1.6948**	0.2990
	2.03	1.78	2.08	0.18
Adjusted R^2	0.71	0.66	0.66	0.82
Sample size	231,282	222,811	7,809	662
Imputed optimal city size	2,086,503	2,208,668	1,686,557	2,257,571
Mean of actual sizes	4,006,440			

Dependent variable: log (total output). Independent variables also include other firm characteristics, two-digit industry fixed effect, dummies for capital cities and cities directly under the central government. Numbers below the coefficients are t statistics. Standard errors are clustered by city-industry combination. "***", "**", and "*" indicate significance at the 1%, 5%, and 10% levels.

4.5 Results by industry

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 to provide additional insight. Out of 30 two-digit manufacturing industries, 28 have positive coefficients of $Log(population)$, 10 are significant at least at the 10% level; 29 have positive coefficient of $Urban\ diversity$, and 18 are significant at least at the 10% level; 21 have positive coefficients of $Specialization$, and 12 are significant at least at the 10% level; 16 have positive coefficients of $Indavedu$, and only one is significant at least at the 10% level; 11 industries show positive coefficients for all the four agglomeration variables. Taking together, the results show that the majority of industries enjoy the benefit from urban diversity and city bigness. There are also some evidence of localization economies. The results are contrast to Nakamura (1985) who found little urbanization economies in manufacturing industries in Japan.

5 CONCLUSION

This paper uses the 2004 China economic census data of manufacturing industries 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, too large a city size (say, over two million population) will generate net diseconomies. In general, firms benefit from urban diversity, but small firms benefit more than large firms do, consistent with Jacobs's idea that small firms rely more on the diverse external environment. We also find that medium cities generate more significant localization economies than large cities do, but large cities generate more significant Jacobs externalities than medium cities do. One policy implication is that while large cities still maintain their advantage of attracting firms of all sizes, medium cities should better attract small and medium firms and encourage clustering and specialization. Another policy implication is that a city should encourage industrial diversity. This is not a conflict to specialization. A city can be relatively diverse while having a few specialized industries, compared with other cities.

We did not address the firm selection issue. It is possible that high-productivity firms may select cities of certain size or cities of certain characteristics, or that small cities with higher productivity grow to large cities at some point of time, creating spurious correlation between productivity and city size or urban diversity. We have used a set of dummy variables to control for unobserved characteristics of firms, industries, and cities, and the conclusions would not be too sensitive to the sorting bias.

REFERENCES

Abdel-Rahman, Hesham M. 2000. "Multi-firm City Versus Company Town: A Microfoundation Model of Localization Economies," *Journal of Regional Science*, 40, 755-769.

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

Arnott, Richard. 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 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, Statistics Canada.

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

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.

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

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

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

Henderson, Vernon. 2003. "Marshall's Scale Economies," *Journal of*

Urban Economics, 53, 1-28.

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.

Hoover, Edgar. 1971. *An Introduction to Regional Economics.* New York: Alfred A. Knopf, INC.

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

Jacobs, Jane. 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 Mion Giordano. 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.

Malizia, Emil and Shanzi Ke. 1993. "The Influence of Economic Diversity on Unemployment and Stability," *Journal of Regional Science*, 33, 221-235.

Moomaw, Ronald. 1981. "Productivity and City Size: A Critique of the Evidence," *Quarterly Journal of Economics*, 96, 675-688.

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.

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

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.

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

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.

Acknowledgement: We thank Richard Arnott, Yuming Fu, the seminar and conference participants at Tsinghua University, Central University of Finance and Economics, the China Economists Society 2008 conference, and the 13th Asian Real Estate Society Annual Meeting for helpful comments.