doi:10.1111/j.1435-5957.2012.00452.x

Contributions of local agglomeration to productivity: Stochastic frontier estimations from Japanese manufacturing firm data*

Ryohei Nakamura[1]

[1] Okayama University, Department of Economics, 3-1-1, Tsushimanaka, Kita-ku, Okayama 800-8530 and Research Institute of Economy, Trade & Industry (RIETI), Kasumigaseki, Chiyoda-ku, Tokyo, Japan
(e-mail: ubbz0252@okayama-u.ac.jp)

Received: 1 July 2011 / Accepted: 14 May 2012

Abstract. This study focuses on local agglomeration effects on productivity, in which the spatial size-distribution of firms is taken into consideration. The data set is comprised of four-digit firm level data on Japanese manufactures for 2005. The estimation by use of firm data enables us to identify various agglomeration effects. The estimation model is based upon the stochastic frontier (labour) production function approach, in which we examine what type of local agglomeration contributes to improve productive efficiency. The estimation results suggest that for most light industries agglomeration of various sizes of firms is important for productivity rather than the concentration of uniform small size firms.

JEL classification: C31, L25, R21

Key words: Local agglomeration, urbanization, stochastic frontier, size of firms

1 Introduction

Several external factors influence firms' productivity. The most important and well-known external factor affecting firm productivity is an agglomeration economy accrued from the spatial concentration of economic activities in a limited area. Various types of firms benefit to various degrees from the agglomeration of economic activity.

The agglomeration economy concept explains modern cities, particularly their productivity, size distribution, and growth. It has policy implications for local municipalities as well as central governments implementing industrial cluster policies.

* This paper is a part of outcomes of the research project 'Sustainable Regional Economic System' conducted at RIETI (Research Institute of Economy, Trade & Industry, Japan) in 2010. An earlier version of this paper was presented at the 51st European Congress of the RSAI in Barcelona. The author thanks to the guest editor Andreas Stephan as well as anonymous referees for helpful comments and suggestions which substantially improved this paper.

In the academic field of agglomeration studies, considerable effort has been devoted to exploring the mechanism of agglomeration and estimating agglomeration effects, and thus a number of empirical studies on agglomeration economies exist. As reviewed by Baldwin et al. (2007), the early stage of agglomeration studies, largely in the 1980s, have focused on the relative strength of urbanization (population or population density) and localization (industry size) in urban productivity using aggregated city or metropolitan area level data.

Urbanization economies are the economic benefits accruing from the geographic concentration of various types of activities in a particular area, reflecting the diversity of urban activities as stated by Jacobs (1969). In contrast, localization economies are benefits yielded by the local concentration of firms in the same industry. According to Marshall (1920), there are three sources of localization economies: input sharing; labour market pooling; and knowledge spillovers.[1]

In the early twentieth century, Marshall found agglomeration sources in several cities in England and describes them in his classical textbook. Recently, the important roles of these sources in local agglomeration have been emphasized by Rosenthal and Strange (2004) and some empirical studies that focus on one source by using micro data (e.g., Overman and Puga 2009).

In the present study, we also focus on local agglomeration, which is the source of localization economies. In investigating local agglomeration phenomena, our fundamental question is, "To what extent do agglomerations of individual firms influence productivity?"

We focus on this issue because most developed countries, particularly Japan, are experiencing a decrease in the number of firms in locally agglomerated areas.[2] There are, of course, production shifts into developing countries due to lower labour costs. In this circumstance, are local agglomeration benefits (localization economies) still effective for productivity?

The contribution of the paper is to analyse the relationship between firm productivity and local agglomeration by firm-level data.[3] The use of firm-level data allows us to estimate several types of effects of the local environment as well as those of city-wide or region-wide agglomerations. For example, urbanization economies usually reflect city-level effects, whereas localization economies tend to occur in more limited areas such as at the post code level, and infrastructure such as highways affect a firm's productivity in a wider area. In addition, as firm-level data provide individual firm size, that is, the number of employees, we can distinguish the effects of large-scale economies from local agglomeration economies. To investigate several agglomeration effects which are different by spatial unit, we construct several indexes in subsequent sections.

In a modern economic theory, the interpretation of Marshall's localization economies is the existence of industry-level economies of scale with constant returns to scale at the individual firm level. In this theory, such economies are derived from the spatial concentration of firms in the same industrial group, particularly the concentration of small- or medium-sized firms. As noted by Holmes and Stevens (2002, 2004), however, the number of firms in a district and the average firm size have a positive correlation, contrary to Marshall's localized agglomeration. In a more recent use of Irish panel data, Barrios et al. (2006) found that firms are relatively larger in districts where industries cluster. Figueiredo et al. (2009), using Portuguese data, found that the relationship between the location quotient, which excludes internal scale effect, and the firm size supports Marshall's local agglomeration with regard to small- and medium-sized firms.

Although the degree of localization and firm size correlate closely (positively), it may depend on the development stage that a nation's industries are in as well as on the industry

[1] A detailed explanation of these sources of agglomeration economies can be found in the review article by Rosenthal and Strange (2004).

[2] According to the survey by Cabinet Office of Japan in recent years, the overseas production ratio for manufacturing industry reached 17.8 per cent in 2009 while it was 6.4 per cent in 1990 and 3.0 per cent in 1985.

[3] A recent agglomeration study applying firm-level data is Anderson and Lööf (2011), which uses Swedish manufacturing data.

itself.[4] To capture Marshall's local agglomeration with greater statistical precision, we need to determine how the spatial distribution of firm size (including distance between firms) affects firms' productivity through the three sources of agglomeration: input sharing; labour market pooling; and knowledge spillovers. We must also acknowledge that agglomeration effects will vary by geographical levels, such as post code and city levels.

We adopt the stochastic frontier production function model which is suitable for the estimation of several types of spatial agglomeration by using micro data. Stochastic frontier approach enables us to distinguish agglomeration effects from firm's technical (in)efficiency. Several earlier studies estimated the stochastic regional production function using spatially aggregate data. However, because the term 'frontier' means individual firms, inefficiency degrees at the regional level, which are estimated by spatially aggregate data, are ambiguous with regard to the region as a whole. In addition, agglomeration is by nature a micro-economic characteristic, making it difficult to distinguish between productive efficiency in terms of stochastic frontier and agglomeration effects using spatially aggregate data.

Our estimations of the stochastic frontier production function with agglomeration use firm-level data comprising four-digit industrial classifications of Japanese manufacturers. Concerning local agglomeration effects, one of our main findings is that the local concentration of various sizes of firms, rather than that of only small-sized firms obtains agglomeration economies of localization. We believe this finding is a significant contribution that focuses on firms' size distribution within a district or region using Japanese manufacturing firm-level data and obtains estimates of agglomeration effects at various spatial levels.

The remainder of this paper is organized as follows. In Section 2 we briefly review of stochastic frontier production function approach by regional (aggregate) data. Section 3 provides a brief review of agglomeration measurements and calculation results for the model specification in the subsequent section. In Section 4, we formulate the size distribution of agglomeration firms within a region to identify the effect of large-sized firms per Marshall's localization, given that the specification of localization externality allows different effects by firm size. In the model we include as contributing factors of productivity: a standard urbanization variable measured by urban population; and infrastructure represented by total highway length in a prefecture. In Section 5, we present estimation results from firm-level data. Section 6 summarizes and concludes the study.

2 Previous studies of stochastic frontier approach in urban/regional production

Few studies adopt a stochastic frontier approach in an urban or regional context for investigating agglomeration economies. To the best of knowledge, only Beeson and Husted (1989) and Puig-Junoy (2001) estimated stochastic frontier production functions at the regional level, and Mitra (1999) and Lall et al. (2004) estimated it using Indian firm-level data. They found more or less the source of differences of regional or firm's efficiency in agglomeration economies as well as other regional specific or firm's specific variables. More recently, Tveteras and Battese (2006) estimated agglomeration economies using the stochastic frontier production function and Otsuka et al. (2010) estimate regional stochastic production function for the Japanese prefecture level which does not correspond to urban data.

In the first two studies, all production functions are formulated at the regional aggregate level with stochastic terms and estimated inefficiencies for regional (state) levels, whereas the third and

[4] For examples, in the 2009 census of economy in Japan, correlation coefficients between the degree of localization in terms of location quotients and the average firm size are 0.645 for electric machinery while 0.346 for furniture and fixtures for all municipalities (Ministry of Internal Affairs and Communications 2011)

fourth papers use efficiency indexes for individual firms. Specifically, Beeson and Husted (1989) first estimated the stochastic frontier production function at the regional (state) level during the period 1959–1972. Their main purposes were to calculate a state's average deviation from an efficient productivity level of a state's manufacturing sector after estimating the translog production function, and then find sources of regional differences in efficiency by regressing the efficiency index on specific regional variables, including agglomeration variables. The major drawback of using aggregate data is the difficulty in identifying agglomeration effects from the degree of efficiency. A recent article by Otsuka et al. (2010) has the same drawback. In addition, they use a prefectural level, which is a substantially wider area than a city or an urban area as the geographical unit. In such a unit, agglomeration economies do not exist and cannot be identified.

The use of firm-level data can differentiate between the local agglomeration effect and productive efficiency because the agglomeration effect is common to firms located in the same city or region, as demonstrated by Tveteras and Battese (2006). Using firm-level data they attempted to separate external agglomeration effects from technical production inefficiency on the basis of the assumption that agglomeration externalities are already embodied in the production frontier, and therefore reduce technical inefficiency. Their empirical evidence came from the salmon aquaculture industry in Norway, which is neither a manufacturing industry nor an example of urban agglomeration.

3 Measuring agglomeration

To specify the estimation model, we briefly summarize indexes of agglomeration measurement and apply them to Japanese manufacturers' firm-level data. It provides important fact findings for specifying agglomeration production function model and also gives the guideline for selecting four-digit industries from the entire industrial classification.

Broadly speaking, two approaches of measuring industrial agglomeration are applied. The first is to measure the spatial distribution of firms across regions within a country by focusing on a particular industry. We acknowledge that the firms in an industry can be uniformly distributed or locally agglomerated. The second is to capture regional specialization, which is usually a relative measure compared to a nation's industrial composition.

3.1 Local agglomeration

We first consider measures of the spatial distribution or industrial localization of industry i in terms of employment, which reflect the geographic concentration of industry i employment across regions. Assuming that there are J regional (or geographical) units and I industries in a country, with the number of industry i employees in region j denoted by x_{ij}, constructing such a measure involves characterizing employment in industry i, region j as a share of employment in industry i across all regions.

One summary measure of geographic concentration based on the employment share measures,

$$s_{ij}^{C} = \frac{x_{ij}}{\sum_{j=1}^{n} x_{ij}} \equiv \frac{x_{ij}}{x_{i*}}, \quad i = 1, \ldots, I\,; \; j = 1, \ldots, J, \tag{1}$$

can be computed as the sum of squares of s_{ij}^{C} across all regions:

$$H_i^C = \sum_{j=1}^{J} \left(s_{ij}^C\right)^2.$$

This is a form of the Hirschman-Herfindahl index (HHI), which is equal to one if the industry is fully concentrated in one region and approaches zero if the industry is evenly distributed with very small shares over a great number of regions.[5]

Another measure of agglomeration combines the spatial concentration measure for industry i in Equation (1) with that for all industries to compute the concentration of industry i in region j relative to that of all industries (or economic size) in region j, compared to the nation as a whole:

$$LQ_{ij}^C = \frac{s_{ij}^C}{s_{*j}} \equiv \frac{x_{ij} / x_{i*}}{x_{*j} / x_{**}}, \quad j = 1, \ldots, J.^6 \tag{2}$$

That is, this form of location quotient (LQ), Equation (2), reflects the percentage (share) of industry i's productive activity in region j relative to the percentage (share) of total productive activity in region j, expressed in terms of employment.

As an alternative to measuring the relative concentration of industry i compared to all industries as a ratio, it can be computed as the difference: $s_{ij}^C - s_{*j}$. If this difference is positive for region j, then region j is more specialized in industry i compared to all industries, or the employment share of industry i in region j is high relative to the share of total employment in region j.

The aggregate of the squared sum of this measure over regions expresses the degree of location deviation or spatial concentration of industry i as:

$$G_i^{C_2} = \frac{1}{J} \sum_{j=1}^{J} \left(s_{ij}^C - s_{*j}\right)^2. \tag{3}$$

These G measures, sometimes called dissimilarity measures, take a value near zero if the spatial distribution of industry i is similar to that of all industries (For example, see Traistaru and Iara 2002).

The value of the index specified in Equation (3) to show industry i's geographic concentration is zero if employment in industry i and total employment have an identical geographic distribution, and continues to increase when each industry is concentrated in few regions. In the case of a few large localized firms, $G_i^{C_2}$ will take a high value for industry i employment without agglomeration of firms' spatial concentration. To identify agglomeration economies implied by such industrial distribution, firm size must be incorporated into this measure.[7]

Ellison and Glaeser (1997) recognize the dependency between industrial distribution and geographic concentration by developing a probabilistic location model based on 'throwing darts' at plants on a country's map. If there are no natural advantages or spillovers between firms, the probability of locating in region j depends solely on the geographical size of the region. However, in the presence of such spatial externalities, agglomeration should be captured by the agglomeration measure.

---

[5] This index was first proposed by Herfindahl in his 1950 Ph.D. dissertation. Hirschman showed its usefulness when he applied the index to industrial concentration.

[6] With regard to asterisks, x_{*j} and x_{i*} denote the corresponding measures of x_{ij} aggregated to the industry and location levels, respectively.

[7] Lafourcade and Mion (2007) investigate the spatial distribution of manufacturing, depending on the size of plants, using Italian data.

Ellison and Glaeser (1997) first define a normalized G measure controlling for the distribution of national employment in industry i, denoting raw geographic concentration:

$$G_i^{(EG)} = \frac{\sum_{j=1}^{J}\left(s_{ij}^{C} - s_{*j}\right)^2}{1 - \sum_{j=1}^{J}\left(s_{*j}\right)^2},$$

where the denominator takes a value of one if total employment in the industry is evenly distributed across regions. On the other hand, plant size distribution in industry i is measured using HHI based on the number of firms, rather than regions or industries:

$$H_i^P = \sum_{k=1}^{K}\left(z_{k\in i}\right)^2, \tag{4}$$

where $z_{k\in i}$ denotes the employment share of firm k in industry i and K is the number of plants in industry i. If all plants have the same size of employment, the inverse of H_i^P collapses to the number of plants, K. The more uneven the firm size distribution, or the higher the level of industrial concentration, the smaller is H_i^P.

Ellison and Glaeser (1997) use the expected value of the raw concentration measure, $E\left[G_i^{(EG)}\right]$, given by

$$E\left[G_i^{(EG)}\right] = \gamma_i\left(1 - H_i^P\right) + H_i^P,$$

to derive the estimator γ_i representing the excess of raw geographic concentration relative to productive concentration with respect to industry i:

$$\hat{\gamma}_i^{(EG)} = \frac{G_i^{(EG)} - H_i^P}{1 - H_i^P} = \frac{\sum_{j=1}^{J}\left(s_{ij}^{C} - s_{*j}\right)^2 - \left(1 - \sum_{j=1}^{J}\left(s_{*j}\right)^2\right)\sum_{k=1}^{K}\left(z_{k\in i}\right)^2}{\left(1 - \sum_{j=1}^{J}\left(s_{*j}\right)^2\right)\left(1 - \sum_{k=1}^{K}\left(z_{k\in i}\right)^2\right)}, \tag{5}$$

where the numerator shows the difference between the degree of geographic concentration of industry i (the spatial HHI) and its expected value (See Ellison and Glaeser 1997, for a detailed explanation and proof). This index, typically called the Ellison and Glaeser index (EGI), indicates whether concentration is greater than the expected value of the random location of firms (without suggesting a reason for the agglomeration). $\hat{\gamma}_i^{(EG)}$ is therefore interpreted as a combined measure of the strength of all agglomeration drivers such as natural advantages and spillovers among firms (see Alecke et al. 2006).

Table 1 shows several measurements of location indexes based upon EGI for all 24 two-digit industries sorted by EG value. EGI is the adjusted value of the raw index by HHI that exhibits the measurement of firms' size distribution at the national level using firm-level data. Industries such as food products with many firms located in many municipalities tend to have lower EGI.

There are several problems with EGI despite its wide usage. First is that the value of the EGI varies with changing geographical units. A wider geographical unit, in general, takes a larger EG value (Rosenthal and Strange, 2001). Second is that sometimes EGI differs from reality by excluding the effect of firms' size distribution. Third, EGI is not a true agglomeration index but a spatial concentration index exhibiting spatial deviation of firms in a given geographical unit.

Table 1. Location deviation: 2-digit level

2- Digit	Industries	Number of municipalities	Number of establishments	Numerator of raw index	Raw Index		HHI of establishment	EGI	
					Rank	Value		Rank	Value
21	Leather tanning, leather products and fur skins	651	6,135	0.01806	1	0.01814	0.00098	1	0.01718
11	Textile mill products	1,019	24,235	0.01306	3	0.01312	0.00038	2	0.01275
30	Transportation equipment	1,393	17,593	0.01286	4	0.01292	0.00243	3	0.01052
15	Pulp, paper and paper products	1,176	11,428	0.00630	7	0.00633	0.00059	4	0.00574
18	Petroleum and coal products	642	1,126	0.01503	2	0.01510	0.00875	5	0.00641
31	Precision instruments and machinery	921	7,451	0.00611	9	0.00614	0.00191	7	0.00424
17	Chemical and allied products	1,048	5,477	0.00509	12	0.00511	0.00107	8	0.00405
28	Information and communication electronics equipment	731	2,831	0.00800	6	0.00803	0.00398	9	0.00407
22	Ceramics, stone and clay products	1,927	21,819	0.00414	13	0.00416	0.00061	11	0.00355
13	Lumber and wood products, except furniture	1,833	16,697	0.00378	15	0.00380	0.00032	12	0.00348
16	Printing and allied industries	1,617	31,970	0.00365	16	0.00366	0.00037	13	0.00330
20	Rubber products	834	5,324	0.00647	8	0.00650	0.00311	14	0.00339
12	Apparel	1,699	29,328	0.00295	20	0.00296	0.00019	16	0.00277
10	Beverages,tobacco and feed	1,417	7,230	0.00400	17	0.00402	0.00087	17	0.00315
19	Plastic products	1,501	23,180	0.00286	21	0.00288	0.00030	19	0.00258
27	Electorical machinery	1,442	16,917	0.00348	18	0.00350	0.00093	20	0.00257
25	Fabricated metal products	1,781	62,849	0.00217	23	0.00218	0.00015	21	0.00203
32	Miscellaneous manufacturing industries	1,666	25,719	0.00319	19	0.00320	0.00121	22	0.00200
26	General machinery	1,659	58,993	0.00223	22	0.00224	0.00038	23	0.00185
9	Food	2,087	48,278	0.00181	24	0.00182	0.00017	24	0.00165

Table 2. Comparison of industries of several indexes

SIC	Firms	Municipalities	Raw index	HHI	EGI
0992 Noodles products	5,202	1,150	0.00515	0.00130	0.00385
2311 Iron industries with blast furnaces	15	14	0.09354	0.09004	0.00385

As an example, we take the case of noodles products (SIC: 0992) and iron industries with blast furnaces. Table 2 displays their primary indicators. Considering the 'raw index' as well as the numbers of firms and municipalities, noodle products' firms seem to be distributed more evenly than those of iron industries. Contrary to intuitive impression, however, both industries' EGI are nearly the same because of the scale effect at the plant level of Iron Industries. Thus, in the subsequent section, instead of EGI for local agglomeration measurement we adopt a combination variable of establishment size and the number of employment by utilizing the benefit of micro data.[8]

3.2 *Intra-regional distribution and agglomeration*

In addition to the national distribution of firms, it is important for agglomeration economies to explore intra-regional distribution of firms as a regional characteristic of agglomeration. The most basic agglomeration measure is the concentration of firms in a limited area, which is captured by the number of firms. However, due to each firm's different employment size, the regional version of HHI based on the number of plants is:

$$H_{ij}^{P} = \sum_{m=1}^{M} \left(z_{m\in i,j} \right)^2, \tag{6}$$

where $z_{m\in i,j}$ denotes the employment share of firm m in industry i, and K is the number of firms in industry i in region j. If all firms have the same size of employment, the inverse of H_i^P collapses to the number of firms, K. The more uneven the firm size distribution, or the higher the level of industrial concentration, the smaller is H_i^P.

Another index for spatial distribution is the entropy measure, which is close to HHI in concept. It is based on the second law of thermodynamics, the entropy law. The entropy index derived from thermodynamics is written as:

$$(Entropy\ Index)_i = -\sum_{j=1}^{m} \frac{x_{ij}}{x_{i*}} \log_2 \frac{x_{ij}}{x_{i*}}. \tag{7}$$

The entropy index takes a value close to zero if most firms in industry i are concentrated in one region, and reaches its maximum value if firms' employment size is uniformly distributed among regions.

The employment share of the largest firms in a region represents the characteristic of firms' size distribution. If the share is large, then one large firm exists. This does not reflect Marshall's local agglomeration in the concentration of small-sized firms.

[8] EGI is not used in the estimation model of stochastic production function as it is the industry index, but it would help the interpretation of estimates by industry.

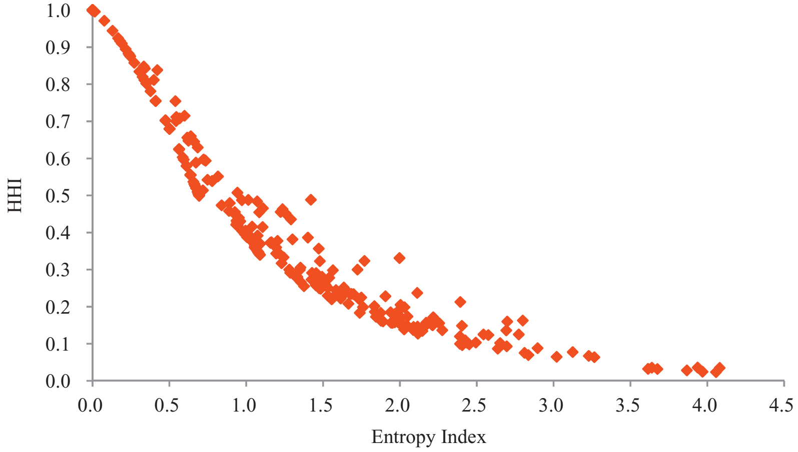

Fig. 1. *Entropy Index* and *HHI*: Wooden furniture

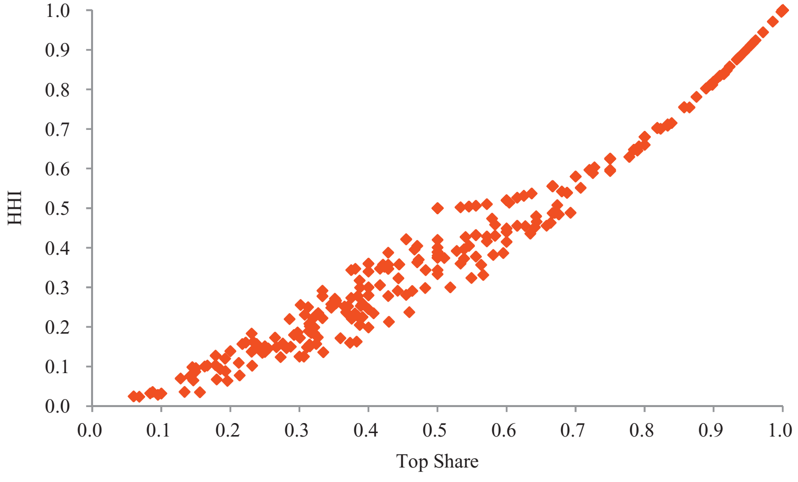

Fig. 2. Share of the Largest Establishment and *HHI*: Wooden furniture

Figure 1 illustrates the relationship between the entropy index and HHI, and Figure 2 shows the share of employees of the largest firm and HHI for wooden furniture (1411). These indexes provide the characteristics of firms' distribution within a region.

The entropy index tends to exhibit a relatively higher value than does HHI in the low level of HHI, which indicates a uniform distribution of activities. From Figure 2 we find nearly a direct positive relationship between the employees' share of the largest firm in a region (denoted as top share) and HHI.

3.3 Regional specialization and agglomeration

We often find regions exhibiting regional specialization in terms of firms' concentration in the same industry in a particular area. In Japan, the following are typical.

Tsubame City in Niigata Prefecture is famous for its agglomeration of tableware, occidental type (SIC: 2521); Imabari City in Ehime Prefecture is famous for its agglomeration of towel products (SIC: 1296); and Sabae City in Fukui Prefecture is famous for its agglomeration of ophthalmic goods, including frames (SIC: 3161). Sabae's ophthalmic goods, including frames account 65.2 per cent of employment and 42.1 per cent of total shipments in Japan. These numbers represent the concentration of small-sized firms in an area. In contrast, tableware in Tsubame in Niigata Prefecture has a 66.7 per cent share of total output in Japan, and the 172 firms constitute 64.2 per cent of Japan's firms in the industry. Imabari City in Ehime Prefecture is famous for its towel production, and its output share is 53 per cent. These areas typify local agglomeration indicating low EG values. We should examine whether these localized firms obtain localization economies.

4 Model and methodology

4.1 Local agglomeration

In Marshall's externality, local agglomeration means the concentration of small-sized firms in a limited area, and economies of localization are originally derived from the local concentration of firms. Thus, it is reasonable that the direct measurement of the localization feature is captured by the number of firms. However, firms' sizes of employment differ, and so HHI is not necessarily small. Local agglomeration can be measured by the number of firms in a limited area if both individual firm sizes and the variance of sizes among firms are small (i.e., HHI is sufficiently small). If there are a few large-sized firms that explain the economies of scale at the plant level, then the magnitude of local agglomeration economies should be reduced even if total employment is large. Therefore, the effect of local agglomeration economies is controlled for plants' or firms' size distribution, although it can be regarded as a function of the number of local firms.

The function of economies of local agglomeration for firm m at location j, $A_{m\in i,j}$, is expressed as:

$$A_{m\in i,j} = g\left(D_{ij'}, l_{i,j}, e_{i,j}\right), \tag{8}$$

where $e_{i,j}$ and $l_{i,j}$ are the number of firms and firms' employees in industry i at post code level j, respectively. $D_{ij'}$ is the variable representing the feature of firms' size distribution within municipality j.[9]

In addition to local agglomeration economies in a limited area, individual firms will have external effects in their productive activities from a wider area than the post code area, such as urbanization economies at the city level or road transportation networks at the prefectural level j. In this study, we adopt the daytime population in city/town $P_{j'}$ as the urbanization variable and total length of highway road $R_{j''}$ as representative of the infrastructure variable, both influencing productivity level as Hicksian neutral factors.[10] By taking account of these external factors, the agglomeration function (8) is rewritten as:

$$A_{m\in i,j} = g\left(D_{ij'}, l_{ij}, e_{ij}, P_{j'}, R_{j''}\right). \tag{8a}$$

[9] Post code level is smaller than regional level.

[10] Infrastructure variables are sometimes treated as unpaid factors of production and atmosphere. In this model, we use road network length at the prefecture level, which can be an external factor to individual establishments. Thus, we define road network length as a Hicks neutral factor that shifts productivity.

The logarithmic specification of local agglomeration (8a) is

$$\ln A_{m\in i,j} = a_0 + a_R \ln R_{j''} + a_P \ln P_{j'} + a_H D_{ij'} + \left(a_l + a_{le} \ln \frac{l_{ij}}{e_{ij}} \right) \ln l_{ij}, \tag{9}$$

where a's are parameters to be estimated. The last term on the right hand side of Equation (9) implies parameter variable specification because the degree of local agglomeration is not only represented by the number of local employees but also depends upon the average firm size as a scale effect. The candidates for $D_{ij'}$ are the entropy index (Equation 7) or HHI (Equation 6) in terms of firms' size distribution, or employees' share of the largest firm in a region. Among these three indexes we adopted HHI in the final estimation model after preliminary regressions.

The elasticity of agglomeration for the number of local firms is derived as:

$$\eta_{m\in i,j} = \frac{\partial \ln A_{m\in i,j}}{\partial \ln l_{ij}} = a_l - a_{le} \ln \frac{l_{ij}}{e_{ij}}, \tag{10}$$

where a_{le} implies the average scale effect of firms on local agglomeration economies.

4.2 Stochastic frontier production function

The nature of local agglomeration effects $A_{m\in i,j}$ is external to individual firms, although certain effects become internalized at the industry level. In our model, it is assumed that the firm-(plant-)level production function is:

$$y_{m\in i,j} = A_{m\in i,j} f\left(k_{m\in i,j}, l_{m\in i,j}\right) \exp\left(v_{m\in i,j}\right), \tag{11}$$

where $y_{m\in i,j}$, $k_{m\in i,j}$, and $l_{m\in i,j}$ are value-added, capital and labour inputs, respectively, for firm m in industry i of region j, and $v_{m\in i,j}$ is a random variable normally distributed with zero mean and constant variance.

In the estimation of Equation (11) we use a more realistic assumption which allows technical inefficiency in individual firms. By considering the firm-specific technical inefficiency term, Equation (11) is rewritten as a stochastic frontier model:

$$y_{m\in i,j} = A_{m\in i,j} f\left(k_{m\in i,j}, l_{m\in i,j}\right) \exp\left(v_{m\in i,j} - u_{m\in i,j}\right), \tag{12}$$

in which $u_{m\in i,j}$ represents a non-negative random variable affiliated with firm-specific factors that would explain the gap between the efficient level and the observed level, and $v_{m\in i,j}$ is a random variable normally distributed with zero mean and constant variance.

The coefficients of the explanatory variables in the inefficiency model (12) are of particular interest to this study. If productive levels, being individual firms' constant primary inputs, vary with agglomeration economies, then agglomeration effects are already reflected in the observed productivity level, whereas technical inefficiency is the difference between the production frontier and the observed output level. Most regional production function models with a stochastic frontier specification have until date included agglomeration effects in the inefficiency term and used agglomeration indexes as explanatory variables for reducing the inefficiency level.[11] However, this is justified only if there are intangible effects of agglomeration in indi-

[11] For example, see Beeson and Husted (1989) and Otsuka et al. (2010) for state and prefecture levels, respectively, aggregated stochastic frontier production function; and Mitra (1999) for plant-level data.

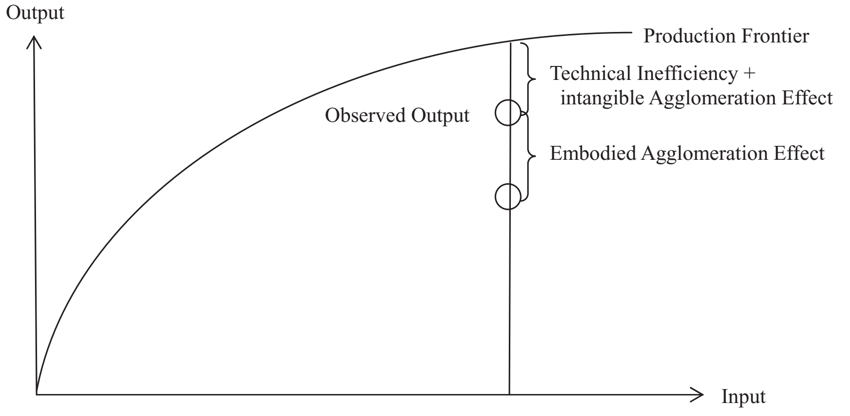

Fig. 3. Agglomeration Effects and Production Frontier

vidual firms' productivity levels (This is depicted in Figure 3).[12] On the basis of this supposition of agglomeration effects, we develop our specifications in subsection 4.3.

4.3 Estimation procedure

Since the work by Aigner et al. (1977) and Battese and Corra (1977), several types of stochastic frontier models for efficiency measurement have been developed, and we can find a comprehensive survey by Kumbhakar and Lovell (2000). Among those models, we adopt the approach proposed by Battese and Coelli (1995) because their single-step simultaneous equation model has an advantage for obtaining efficient parameter estimates. However, we cannot directly apply their approach since our dataset is cross-section, not panel. Some specific distributional assumptions about the one-sided component of disturbance term should be made to obtain efficient estimates. Following the cross-sectional model of Huang and Liu (1994) and the survey by Kumbhakar and Lovell (2000), we assume that errors representing inefficiency follow half-normal distribution. This corresponds to the time-invariant technical inefficiency case of panel data.

Considering the specification (9), the Cobb-Douglas specification of Equation (12) is written as:[13]

$$\ln y_{m\in i,j} = a_0 + a_{HR}\ln R_{j'} + a_P \ln P_j + a_H HHI_{ij'} + \left(a_l + a_{le}\ln\frac{l_{ij}}{e_{ij}}\right)\ln l_{ij} + a_K \ln k_{m\in i,j} + a_L \ln l_{m\in i,j} + v_{m\in i,j} - u_{m\in i,j}.$$

This stochastic frontier production function is not restricted to constant returns to scale at the firm level. To avoid multicollinearity between explanatory variables, we estimate the labour productivity function as:[14]

[12] Similar statements are found in Tveteras and Battese (2006).

[13] In the subsequent implementation of estimation we attempted flexible functional forms such as translog. However, estimated parameters were suffered from multicollinearity because of multiplicative and cross terms. Thus, we adopt the Cobb-Douglas specification.

[14] Baldwin et al. (2010) estimate the labour productivity production function for identifying agglomeration economies with micro data.

$$\ln \frac{y_{m\in i,j}}{l_{m\in i,j}} = a_0 + a_{HR} \ln R_{j''} + a_P \ln P_{j'} + a_H HHI_{ij'} + \left(a_l + a_{le} \ln \frac{l_{ij}}{e_{ij}} \right) \ln l_{ij} + a_K \ln \frac{k_{m\in i,j}}{l_{m\in i,j}} + (a_K + a_L - 1) \ln l_{m\in i,j} + v_{m\in i,j} - u_{m\in i,j}, \tag{13}$$

with the assumption that the errors representing the inefficiency $u_{m\in i,j}$ follow truncated half-normal distribution, where $u_{m\in i,j}$ is a function of observable explanatory variables and unknown coefficients.[15] In Equation (13) technical inefficiency effects are assumed to be defined by:

$$u_{m\in i,j} = b_0 + b_F \ln F_{m\in i,j} + b_{SV} \frac{y_{m\in i,j}}{q_{m\in i,j}} + b_P \ln P_{j'} + b_H HHI_{ij'} + b_L \ln l_{m\in i,j} + w_{m\in i,j}, \tag{14}$$ [16]

where $F_{m\in i,j}$ stands for the amount of the stated capital implying firm's financial muscle. The expected sign of parameter b_F is negative, since firms with larger financial resources can potentially afford to hire well-educated labour and physical capital with higher quality. Another explanatory variable for unobserved firm's inefficiency is value-added ratio defined as $y_{m\in i,j}/q_{m\in i,j}$, where $q_{m\in i,j}$ is shipment as output measurement. Firms exhibiting high value-added ratio in an industry are expected to operate plants more efficiently. Both variables are expected to reduce inefficiency. We also add agglomeration variables in Equation (14). As in Figure 3, if agglomeration externalities are not fully manifested in the realized output level, then inefficiency term includes (local) agglomeration measurements as shift variables. To examine those effects three (local) agglomeration indexes, which are $P_{j'}$, $HHI_{ij'}$, and $l_{m\in i,z\in j}$ are incorporated in Equation (14). $w_{m\in i,j}$ is a random variable defined as the truncation of the normal distribution with zero mean and variance σ^2.

Parameter estimates of Equations (13) and (14) are obtained by applying the maximum likelihood method following Aigner et al. (1977), Battese and Broca (1997) together with the variance parameters expressed as:

$$\tilde{\sigma} = \sqrt{\sigma_v^2 + \sigma^2} \quad and \quad \gamma = \sigma / \sigma_v,$$

where σ is the standard deviation of the $N(0, \sigma^2)$ distribution required for non-negative errors, $u_{m\in i,j}$, and σ_v is the standard deviation of symmetric errors $v_{m\in i,j}$.

5 Empirical analysis

5.1 Data and variables

The primary data base of this study is individual firms drawn from the 2005 census of manufactures in Japan. The census survey is conducted by METI (Ministry of Economy, Trade, and Industry 2007) annually; years whose last digit is 3, 5, 8, or 0, the survey is conducted on the entire sample, whereas for other years the census covers only firms with four or more employees. In this study the most recent available data was for 2005.

Value added is defined as gross value-added, which is the value of manufactured goods shipments minus the value of raw materials, fuels and electricity consumed, and expenses for

[15] For example, Battese and Broca (1997) adopt the assumption of half-normal distribution.
[16] As dependent variable we use non-negative $u_{m\in i,j}$ of each firm like Mitra (1999), while Tveteras and Battese (2006) adopt distributional mean of $u_{m\in i,j}$.

consigned production. Labour is defined as the number of employees, including part-time as well as full-time workers. Capital stock is the value of tangible fixed assets, including land and physical equipment at the start of the year. Unfortunately, the survey for capital assets includes only firms with 10 or more employees, but many firms have fewer than 10 employees. Concerning wood furniture (SIC: 1411), for example, of the total 8,029 firms, 1,236 firms have under 10 employees. To avoid reducing samples in the estimation model, we add a dummy variable on the capital variable parameter for small-sized firms. Thus, the number of samples in the estimation remains the whole sample.

These firm- (plant-)level data are combined with several municipal-level data such as daytime population, land area available for living, total length of highway.[17] The analysis is performed for 17 industrial sectors in three categories. The criteria for categorizing four-digit industries are as follows:

- traditional and local manufacturing industries;
- industries requiring skilled labour; and
- high-tech industries

In 2005, the number of municipalities was 2,163, and there were 590 four-digit industries.

5.1.1 Traditional and local manufacturing industries

We choose six typical industries that are traditional or local, some of which are declining due to the increase of imported goods from developing Asian countries. Japanese sake products (1023) and wooden furniture are traditional industries distributed over many municipalities, and thus their EGIs are low. In contrast, the industries of towel products, plastic shoes, and clay roofing tile are locating in specific municipalities, and thus their EGIs are high (see Table 3).[18]

Table 3. Location characteristics of traditional and local manufacturing industries

	0923 Fish paste	1023 Sake (rice wine)	1141 Fabric mills	1212 Ladies garments
EGI	0.00703	0.00815	0.01917	0.00529
	(284)	(443)	(164)	(317)
HHI	0.00474	0.00316	0.00331	0.00667
	(466)	(500)	(497)	(435)
Municipalities/firms	404/1471	692/2,905	219/2,023	1,046/6,927
	1296 Towel	1411 Wooden furniture	2022 Plastic shoes	2231 Clay roofing tile
EGI	0.26896	0.00490	0.13032	0.07884
	(7)	(324)	(17)	(36)
HHI	0.00826	0.00093	0.03662	0.00750
	(404)	(550)	(185)	(417)
Municipalities/firms	66/461	1,184/8,030	90/792	139/576

Note: Numbers in parentheses are descending in order.

[17] Data characteristics are presented in the Appendix.
[18] Kobe accounts for 14.4 per cent of the Japanese Sake output, but only 1.2 per cent of the number of establishments. This suggests that a few large establishments are located in Kobe. However, EGI will not have a high value because of firms' scale adjustment, which is one of its drawbacks.

5.1.2 Industries requiring skilled labour

We select four representative industries at the four-digit level: moulds and dies, parts, bicycles and parts, ophthalmic goods, frames, and jewellery products (Table 4). Among them, ophthalmic goods, frames is also regarded as a traditional and local industry. Firms belonging to the moulds and dies, parts' industry are located in many municipalities, and thus the industry exhibits a low EGI. The moulds and dies, parts' industry is deemed to be closely related to machinery production or metal products in its input-output processes. Thus, those industries may exhibit co-agglomeration.

Table 4. Location characteristics of industries required skilled labour

	2696 Moulds and dies, parts	3091 Bicycles and parts	3161 Ophthalmic goods, frames	3211 Jewellery products
EGI	0.00322	0.10720	0.23802	0.03543
	(542)	(26)	(9)	(90)
HHI	0.00120	0.03835	0.01278	0.00671
	(542)	(178)	(351)	(433)
Municipalities/firms	978/9,984	133/371	71/804	254/1,166

Note: Numbers in parentheses are descending in order.

5.1.3 High-tech industries

It is important for industrial policy to address whether high-tech industries obtain local agglomeration economies. We select three industries at the three-digit level and four industries at the four-digit level (Table 5). EGIs of these industries are low although there are few firms. In particular, HHI of 'print circuits' is markedly low. Thus, these industries are not very agglomerated in specific areas compared to traditional industries.

Table 5. Location characteristics of high-tech industries

	2810 Communication equipment	2820 Electronic computer	2910 Electronic parts and devices
EGI	0.00498	0.00372	0.00368
	(77)	(90)	(92)
HHI	0.00686	0.00930	0.00162
	(63)	(53)	(119)
Municipalities/firms	550/1,665	467/1,166	1,213/7,408

	2812 Communication equipment	2822 Personal computer	2912 Semiconductor	2913 Integrated circuit	2918 Print circuit
EGI	0.00492	0.00594	0.00350	0.00496	0.00474
	(387)	(341)	(457)	(385)	(399)
HHI	0.11938	0.01979	0.02371	0.01360	0.00428
	(240)	(275)	(247)	(342)	(474)
Municipalities/firms	220/398	214/366	137/173	145/197	563/1,610

Note: Numbers in parentheses are descending in order.

5.2 *Estimation results and interpretations*

We estimate the stochastic frontier production function and inefficiency function using the maximum likelihood method. Table 6 reports the estimated parameters and certain indicators representing industry characteristics.[19] In the stochastic frontier production function model, several external factors incorporated as explanatory variables are the different geographical units.

5.2.1 Highway effect

The variable most widely affecting individual firms as an external effect is the highway network, measured by the total length of highway as transportation infrastructure at the prefecture level ($R_{j'}$).[20] Highway road lengths vary considerably among prefectures while the average over prefectures is 178.9 km. In general, transport infrastructure without congestion has a positive effect on manufacturing. As expected, most of the industries show positive signs, some with significant values. Among them, the following industries that use trucks for transportation of their final goods have relatively higher estimates and significant t-values at 1 per cent level; fish paste products (0923), fabric mills and wooden cotton (1141), wooden furniture (1411), and jewellery products (3211).

5.2.2 Urbanization effect

In general, urbanization externally affects firms' productivity at the urban/city level because urbanized areas form cities. In this study, estimated parameters of urbanization economies, a_P, have positive and significant values in most of the industries because all the industries chosen for this estimation belong to the light industry category, which enjoys urbanization economies.[21] In particular, ladies garments (1212), wooden furniture (1411), plastic shoes (2022), and high-tech industries obtain high economies of urbanization. These industries tend to be located in urbanized areas, whereas industries with less urbanization economies, such as clay roofing tiles (2231) or towel products (1296), are located in somewhat less urbanized areas, as they belong to traditional and local manufacturing industries. The values of the estimated parameter imply a percentage productivity increase when the city's size doubles. For example, for plastic shoes (2022), the estimated parameter 0.0636 means a 6.36 per cent increase in productivity if the city size, represented by daytime population, doubles.

5.2.3 Firm-size distribution effect

We adopt HHI to capture the firms' size distribution within a region. HHIs are calculated at the municipal level. If there are a few large firms in a concentrated area, then the HHI takes a large value compared to the situation with evenly distributed firms. Thus, the negative sign of the

[19] To check robustness, it may be necessary to compare another approach such as data envelope analysis (DEA) or total factor productivity (TFP). DEA is a different approach from stochastic frontier approach although both treat inefficiency. TFP is generally used in time series or panel analysis; however, TFP measurement is not suitable in our cross-section estimation. Panel data is now under construction for future research.

[20] Prefecture is an independent local governmental unit which is similar to a state or province. There are 47 prefectures in Japan.

[21] For example, see Nakamura's (1985) results for Japan.

Table 6. Estimated results of stochastic frontier production function and efficiency function (traditional and local manufacturing industries)

Variable	Parameter	0923 Fish paste products	1023 Sake (rice wine)	1141 Fabric mills woven cotton	1212 Ladies garments
Constant	a_0	5.0384**	5.1218**	5.8866**	4.3386**
		(13.17)	(15.08)	(12.01)	(19.94)
Highway road length	a_{HR}	0.0433**	0.0191	0.0702**	0.0244
		(3.25)	(1.75)	(4.31)	(1.55)
Daytime population	a_P	0.0287	0.0592*	–0.0625**	0.0571**
		(1.57)	(2.36)	(–3.83)	(6.42)
HHI	a_H	0.0525	0.0056	–0.1006*	0.0338
		(1.73)	(1.59)	(–2.44)	(1.33)
Number of workers at post code	a_l	0.0836*	0.2021**	0.0141	0.0625**
		(1.98)	(2.81)	(1.10)	(4.80)
Average size of firm	a_{le}	0.0364**	0.0559**	0.0454**	0.0151*
		(4.41)	(4.28)	(3.12)	(2.33)
Capital stock	a_K	0.3455**	0.4853**	0.1781**	0.1816**
		(7.99)	(9.84)	(3.61)	(7.67)
Number of workers	a_L	0.6058**	0.4989**	0.8420**	0.8298**
		(13.21)	(10.31)	(23.86)	(13.67)
Scale economies		–0.0487	–0.0158	0.0201	0.0114
$1 - \alpha_K - \alpha_L$		(–0.55)	(–0.32)	(1.57)	(1.85)
Average technical efficiency		0.6251	0.6388	0.7593	0.7381
Average elasticity of local agglomeration	$\bar{\eta}_{i,j}$	0.0169	0.0261	0.0398	0.0453
Inefficiency model					
Constant	b_0	4.4421**	3.7544**	11.0127**	6.0581**
		(9.01)	(10.14)	(3.88)	(3.04)
Amount of stated capital	b_F	–0.3039**	–0.2327**	–0.3916**	–1.8300
		(–5.04)	(–4.99)	(–2.71)	(–1.25)
Value-added ratio	b_{SV}	–0.4873**	–0.5886**	–0.1211*	–0.5369*
		(–4.76)	(–5.18)	(–2.49)	(–2.15)
Daytime population	b_P	–0.0052	–0.0225**	–00697**	–0.0869*
		(–0.99)	(–2.88)	(2.79)	(–2.15)
HHI	b_H	–0.0794	0.1332	0.2916*	–0.4369*
		(–0.29)	(1.67)	(2.06)	(–2.35)
Number of workers at post code	b_{ZL}	–0.0938	–0.6437**	–0.5520	–0.5369*
		(–0.56)	(–5.58)	(–1.69)	(–2.15)
	γ	0.7932**	0.8627**	0.8440**	0.8960**
		(14.11)	(25.65)	(23.66)	(9.34)
	σ^2	1.0976**	1.4394**	3.1616**	2.6385**
		(4.68)	(4.35)	(9.25)	(8.57)
Log likelihood		–1,611.93	–1,618.21	–2,534.47	3,127.13
Observations		1,457	1,466	2,014	6,911

Table 6. Continued

	Parameter	1296 Towel products	1411 Wooden furniture	2022 Plastic shoes	2231 Clay roofing tile
Constant	a_0	7.3126**	5.8715**	4.6618**	4.0455**
		(2.87)	(32.16)	(5.23)	(4.87)
Highway road length	a_{HR}	0.0375*	0.0531**	–0.0300	0.0578
		(2.32)	(4.33)	(–0.47)	(1.88)
Daytime population	a_P	–0.0185	0.0485**	0.0636**	–0.0377
		(–0.79)	(4.59)	(5.89)	(–1.22)
HHI	a_H	–0.3087**	–0.0403*	–0.0655**	–0.0245
		(–2.89)	(–2.25)	(–2.81)	(–0.78)
Number of workers at post code	a_l	0.0359**	0.0155*	0.0607*	0.0302
		(2.58)	(2.39)	(2.22)	(1.92)
Average size of firm	a_{le}	0.0353**	0.0450**	0.0065*	0.0499**
		(2.63)	(4.31)	(2.37)	(2.83)
Capital stock	a_K	0.2794**	0.1927**	0.1485*	0.2667**
		(3.34)	(6.81)	(2.50)	(3.83)
Number of workers	a_L	0.7574**	0.8543**	0.8710**	0.8105**
		(11.49)	(46.24)	(16.81)	(18.80)
Scale economies		0.0368	0.0470	0.0195	0.0772
$1 - \alpha_K - \alpha_L$		(1.33)	(1.72)	(1.77)	(1.95)
Average technical efficiency		0.7031	0.7387	0.6934	0.6567
Average elasticity of local agglomeration	$\bar{\eta}_{i,j}$	0.0637	0.0601	0.1234	0.0545
Inefficiency model					
Constant	b_0	4.1142**	5.5958**	1.5054*	4.0687**
		(11.00)	(7.32)	(2.36)	(13.40)
Amount of Stated capital	b_F	–0.2809**	–1.26980**	–0.3129**	–0.2345**
		(–7.62)	(–5.07)	(–4.46)	(–5.07)
Value-added ratio	b_{SV}	–0.7124**	–0.5469**	–0.0245	–0.3003**
		(–3.99)	(–4.75)	(–1.58)	(–5.07)
Daytime population	b_P	–0.0859*	0.0254	0.0520	–0.2017*
		(–2.41)	(0.88)	(1.33)	(–2.25)
HHI	b_H	–0.1665**	0.1290*	0.1587	–0.5369*
		(–6.14)	(2.39)	(1.89)	(–2.15)
Number of workers at post code	b_{ZL}	–0.2068	–0.4070**	–0.2005*	–0.2509*
		(–1.37)	(–3.11)	(2.59)	(–2.09)
	γ	0.8489**	0.9028**	0.6144**	0.9079**
		(9.27)	(23.11)	(6.44)	(8.79)
	σ^2	1.7638**	3.24272**	1.9965**	1.7508**
		(9.53)	(5.45)	(5.67)	(6.47)
Log likelihood		–571.10	–3,351.02	–1,072.1	–659.11
Observations		459	7,952	787	566

Table 6. Continued (skilled labour required industries)

	Parameter	2696 Moulds and dies, parts	3091 Bicycles and parts	3161 Ophthalmic goods and frames	3211 Jewellery products
Constant	a_0	5.9801**	4.2369**	6.4628**	5.7796**
		(40.10)	(4.06)	(8.97)	(6.72)
Highway road length	a_{HR}	0.0394*	–0.0223	–0.0233	0.0653**
		(2.15)	(–1.76)	(–1.72)	(3.34)
Daytime population	a_P	0.0215**	0.0276	0.0383*	0.0140
		(3.14)	(1.77)	(2.21)	(1.28)
HHI	a_H	–0.0705*	–0.1278**	–0.0898*	–0.1233**
		(–2.36)	(–2.73)	(–2.11)	(–3.20)
Number of workers at post code	a_l	0.0701**	0.0890*	0.0257	0.1431**
		(6.55)	(2.25)	(1.63)	(3.60)
Average size of firm	a_{le}	–0.0561**	–0.1044	–0.0335	0.0076
		(–4.47)	(–1.75)	(–1.64)	(0.86)
Capital stock	a_K	0.3023**	0.3901**	0.1642*	0.1772*
		(9.30)	(7.63)	(2.50)	(2.51)
Number of workers	a_L	0.6741**	0.6105**	0.8767**	0.8570**
		(29.04)	(10.63)	(16.62)	(19.41)
Scale economies		–0.0236	0.0001	0.0409	0.0342
$1 - \alpha_K - \alpha_L$		(–1.45)	(0.50)	(1.24)	(1.67)
Average technical efficiency		0.7417	0.7789	0.7278	0.6225
Average elasticity of local agglomeration	$\bar{\eta}_{i,j}$	0.0797	0.1798	0.0270	0.0583
Inefficiency model					
Constant	b_0	7.8328**	4.9050**	6.4315*	5.8101**
		(14.81)	(3.19)	(2.17)	(5.28)
Amount of stated capital	b_F	–0.1374**	–0.1047**	–0.2999	–0.3271**
		(–7.51)	(–3.49)	(–1.57)	(–4.76)
Value-added ratio	b_{SV}	–0.8038**	–0.4899**	–0.4199**	–0.1067**
		(–7.39)	(–2.82)	(–2.88)	(–4.34)
Daytime population	b_P	–0.4948**	0.0531	–0.1213*	–0.1033
		(–5.06)	(0.73)	(1.99)	(–1.79)
HHI	b_H	–0.1186	0.0435	0.1574	0.1233*
		(–1.56)	(1.08)	(1.54)	(2.10)
Number of workers at post code	b_{ZL}	–0.4579**	–0.2875*	0.1081	–0.5030**
		(–4.77)	(–2.08)	(1.40)	(–2.85)
	γ	0.9621**	0.7535**	0.9552	0.8703**
		(16.94)	(6.93)	(12.21)	(3.47)
	σ^2	3.6610**	1.6945**	1.7051**	1.1713**
		(9.40)	(5.32)	(8.04)	(5.18)
Log likelihood		–4,012.11	–464.28	–909.56	–1,563.94
Observations		9,943	369	796	1,154

Table 6. Continued (high-tech industries)

	Parameter	2810 Communication equipment	2820 Electronic computer	2910 Electronic parts and devices
Constant	a_0	4.6423**	4.3122**	4.1577**
		(12.66)	(6.22)	(23.34)
Highway road length	a_{HR}	0.0408	0.0605*	0.0394*
		(1.49)	(2.40)	(2.18)
Daytime population	a_P	0.0680**	0.0479*	0.0700**
		(3.98)	(2.45)	(8.01)
HHI	a_H	0.1059	0.1149*	0.1243**
		(1.37)	(2.11)	(3.29)
Number of workers at post code	a_l	0.1396**	0.0519**	0.0430**
		(3.69)	(5.67)	(8.45)
Average size of firm	a_{le}	0.0256**	0.0266**	0.0195**
		(3.70)	(2.93)	(5.84)
Capital stock	a_K	0.3614**	0.3600**	0.3665**
		(21.17)	(10.60)	(30.55)
Number of workers	a_L	0.6885**	0.7324**	0.6944**
		(26.03)	(13.40)	(31.31)
Scale economies		0.0499	0.0924	0.0599
$1 - \alpha_K - \alpha_L$		(1.91)	(1.95)	(1.87)
Average technical efficiency		0.7243	0.6715	0.6988
Average elasticity of local agglomeration	$\bar{\eta}_{i,j}$	0.0465	0.0582	0.0697
Inefficiency model				
Constant	b_0	4.0018**	2.6874**	4.9688**
		(2.67)	(2.87)	(28.77)
Amount of stated capital	b_F	–0.5692**	–0.2848**	–0.7036**
		(–3.08)	(–3.88)	(–16.09)
Value-added ratio	b_{SV}	–0.2065**	–0.1478	–0.2979**
		(–5.14)	(–1.88)	(–16.40)
Daytime population	b_P	0.1550	–0.2051	–0.3369*
		(1.52)	(–1.50)	(–2.15)
HHI	b_H	–0.0658	–0.1421	–0.2367*
		(–1.38)	(–1.09)	(–2.05)
Number of workers at post code	b_{ZL}	–0.3231*	–0.1203	–0.4103*
		(–2.25)	(–1.88)	(–2.17)
	γ	0.8279**	0.7267**	0.7786**
		(19.12)	(5.08)	(19.96)
	σ^2	2.3287**	1.0481**	1.8814**
		(5.79)	(6.18)	(5.90)
Log likelihood		–2,006.08	–1,399.60	–8,120.89
Observations		1,649	1,160	7,337

Notes: All variables except HHI are natural logarithms.
Numbers in parentheses below estimated values are *t*-values.
** Significance at 1% level.
* Significance at 5% level.

estimated parameter means a positive effect of the local agglomeration of small- or medium-sized firms per Marshall (1920). Table 6, however, reports that all estimated parameters for high-tech industries (communication equipment, electronic computer, and electronic devices) are positive and significant; thus, the existence of large factory firms may affect positive externalities for nearby small-sized firms. Japan does, in fact, have agglomerated areas of high-tech industrial firms with one large core factory surrounded by related small firms.

In contrast, for towel products, bicycles and parts, moulds and dies, parts, ophthalmic goods and frames, and jewellery products, the existence of a large-sized firm yields negative externality for other small-sized factories. In particular, towel products and bicycles and parts show 1 per cent significant and 5 per cent significance level and higher negative values, respectively greater than 0.12. These industries all belong to those requiring skilled labour, and we often find that these are skilled-labour intensive, and small-sized firms tend to agglomerate in a limited area.

5.2.4 Local agglomeration effect

Local agglomeration effects can be investigated by the estimated parameters of a_l and a_{le}. a_l captures the direct effect of the concentration of firms measured by the number of employees, and a_{le} is the coefficient for the average firm size variable, and therefore represents the conditional benefit of the local firm size effect, given the local concentration of firms. Because these variables are measured at the seven-digit post code level representing local agglomeration, we observe the neighbourhood effect of localization. The direct effect of local agglomeration measured by the number of local workers shows positive signs for nearly all industries chosen. In particular, bicycle and parts and Japanese sake products indicate significantly high values at 1 per cent and 5 per cent levels respectively. For average firm size effect in an agglomerated area, all traditional and local industries have significantly positive estimated values, whereas industries requiring skilled labour have negative parameters. That the positive parameters of estimated values of a_{le} enhance the local agglomeration effect by the large average firm size, suggests that the existence of large-sized firms may affect possible positive externality for small-sized firms, unlike Marshall's local agglomeration concept.

Equation (5) calculates the elasticity of local agglomeration related to the number of firms. In Table 6, the average elasticity by industry is reported for the entire sample. The highest elasticity is 0.1898 for bicycle and parts (3091). There are, in fact, many small-sized firms in Sakai City in Osaka Prefecture. Plastic shoes (2022) also exhibits a relatively high value of elasticity of local agglomeration, 0.1234. In addition to bicycle and parts (3091), moulds and dies, parts (2692) has a high value. These industries need skilled labour and are usually small in size and localized in Marshall's sense, which means they obtain high benefits of local agglomeration. Firms in the electric parts and devices (2910) manufacturing industry also obtain strong benefits from local agglomeration. In the four-digit industrial classification electric parts and devices includes semi-conductors, printed circuits, and devices, whereas we estimated these at the three-digit level.

5.2.5 Returns to scale

In the model estimation, we do not impose the restriction of constant returns to scale at the firm level. The levels of returns to scale are obtained indirectly from the parameters of labour share (a_K) and capital share (a_L). The results are also reported in Table 6. Most of the industries exhibit weak economies of scale at the firm level.

5.2.6 Technical inefficiency

Table 6 also reports technical efficiencies averaged across sample observations by industry. The efficiency levels vary from 0.60 to 0.75. Inefficiency model (14) includes two firm-specific explanatory variables and three agglomeration variables. Firm-specific variables are the amount of stated capital and ratio of value-added to shipment. The amount of stated capital implies firm's potential financial strength.[22] The larger finance has possibility reduce inefficiency. Firms exhibiting high valued-added ratio may operate plants more efficiently. Agglomeration variables are daytime population as urbanization measurement, HHI as firm's size distribution, and local employees at post code level as local agglomeration. If agglomeration economies are not fully embodied in production process, then intangible effects of agglomeration may be reflected in technical inefficiencies.

All industries exhibit negative and significant t-values for the estimated parameters of capital fund size, and nearly all the industries have negative values for the value-added ratio. These variables substantially reduce the inefficiency of firm productivity as $u_{m\in i,j}$ exhibits the degree of inefficiency.

In Table 7 we present characteristics of agglomeration effects by industry in order to compare embodied affects and intangible effects reflected in production function and inefficiency function, respectively. With regard to urbanization economies, while most of the

Table 7. Signs of estimated parameters for production and inefficiency functions

Industry	Urbanization		Local distribution		Local agglomeration	
	A	B	A	B	A	B
Traditional and local manufacturing industries						
Fish paste products	*	-	*	-	**	-
Sake (rice wine)	**	—	*	*	***	—
Fabric mills, woven cotton	—	—	–	-	*	-
Ladies garments	***	–	*	–	***	–
Towel products	-	—	—	—	***	-
Wooden furniture	***	*	–	**	**	—
Plastic shoes	***	*	—	*	**	–
Clay roofing tile	-	–	-	–	*	–
Skilled labour required industries						
Moulds and dies, parts	***	—	–	-	***	—
Bicycles and parts	*	*	—	*	**	–
Ophthalmic goods and frames	**	–	–	*	*	*
Jewellery products	*	-	—	**	***	—
High-tech industries						
Communication equipment	***	-	*	-	***	–
Electronic computer	**	-	**	-	***	-
Electronic parts and devices	***	–	***	–	***	–

Notes: Urbanization: daytime population at city level. Local distribution: HHI within a city. Local agglomeration: number of employees at post code level.
A: sign of estimated parameter for production function, B: sign of estimated parameter for inefficiency function.
*: positive, **: positive and significant at 5% level, ***: positive and significant at 1% level,
-: negative, –: negative and significant at 5% level, —: negative and significant at 1% level.

[22] The amount of stated capital is different from tangible capital stock. It literally means financial capital that firm can reserve. As shown in the Appendix this variable is not strongly correlated to value-added compared to tangible capital stock.

industries receive external benefits from larger daytime population, some industries which show negative and significant signs for inefficiency function can afford to find benefits of urbanization economies. Those industries are sake (rice wine), moulds and dies, parts, ophthalmic goods and frames, jewellery products, and high-tech industries. On the contrary, fabric mills, woven cotton only shows negative and significant signs of urbanization parameters for both functions. As this industry is a traditional but declining industry in Japan, urbanization effects may not be effective.

Local distribution of firms is measured by HHI. The negative sign of coefficient of HHI means the existence of Marshall's externality because the small number of HHI shows the small variance of firm sizes and relatively small firms' concentration. For industries exhibiting negative signs in column A in Table 7 agglomeration effects are already embodied in production activities, and those exhibiting positive signs in column B can potentially realize local agglomeration benefits. For towel product agglomeration effects are fully embodied in the production activity, since both signs (columns A and B) are negative and significant. Conversely, for high-tech industries, such as electric parts and devices, receive benefits from the existence of large size firms (positive sign of column A and negative sign of column B).

When we look at agglomeration effects more locally (last two columns in Table 7), which means at seven-digit postal code area, most of the industries are able to realize more local agglomeration benefits while those effects are already embodied in the production activity. These fact findings state that local and traditional industries, which are generally declining in Japan, have still possibility to utilize local agglomeration for technological development leading to produce higher value-added goods with high labour productivity.

In summary, we provide the estimation results for the three industrial categories examined. For industries in the high-tech industrial category, the existence of a large-sized firm in addition to the local agglomeration effect positively influences labour productivity. Urbanization economies have positively significant effects on industries in urban areas. In contrast, the industries requiring skilled labour, such as jewellery products, obtain benefit from local agglomeration with small- or medium-sized firms and negative benefits from the existence of a large-sized firm in the area. Traditional and local industries often have unusual, low-tech technologies, such as those in the sake (rice wine) and towel products industries, and benefit from local agglomeration with small firms according to the negative parameters of their HHI.

6 Summary and conclusion

In this study, we estimated local agglomeration economies with diverse variables in different geographical units by applying a stochastic frontier model at the firm level. As agglomeration economies have micro characteristics by nature, it is difficult to distinguish among several externalities influencing productivity from spatially aggregate data (state level or prefecture level) corresponding to NUTS 2 in Europe. Our micro firm data enabled us to obtain the effects of externalities. In the estimation, we chose industries of traditional and local manufacturing, industries with highly skilled labour, and high-tech industries. From the estimation results, we obtained several findings about local agglomeration.

First, the role of public transportation infrastructure represented by highway length is important to most industries. Total highway length at the prefectural level generally contributes to manufacturing firms' productivity, particularly for industries using trucks for transportation of final goods. Second, the daytime population size at the municipal level contributes positively to productivity in terms of urbanization economies in most light industries, particularly high-tech industries. Third, the existence of large-sized firms in an agglomerated area sometimes negatively affects locally agglomerated firms, particularly in industries requiring skilled labour. In contrast, large-sized firms provide external benefits for high-tech industries.

Recently, in locally agglomerated areas, traditional and local industries have been declining, industries requiring skilled labour lack followers, and high-tech industries are seeking to produce new value-added commodities. For policy purposes, we must determine the appropriate agglomeration of firms to benefit from several externalities. On the basis of our estimated results by industrial category, we must consider local agglomeration to achieve that goal.

Our estimated results suggest that for productivity in most light industries, agglomeration of various sizes of firms is more important than the concentration of uniform small-sized firms. However, certain exceptional industries such as bicycle and parts (3091) or moulds and dies, parts (2969) enjoy mutual externalities from the concentration of small-sized firms. For traditional and local industries, policies promoting mutual external effects are needed, whereas areas with agglomerated high-tech industries need policies stimulating collaboration between large- and small-sized firms. Local agglomeration has still potential to enhance labour productivity for local and traditional industries. In order to realize it, several kinds of collaborations among firms at locally agglomerated areas will be necessary as well as governmental supports.

For future studies, several aspects must be considered. First, as we perceived economies of scale at the firm level from estimated parameters of production function, the monopolistic competition model incorporating intermediate goods should be adopted. Second, to obtain more robust estimation results, we need to construct a panel dataset. Finally, co-agglomeration effects, meaning forward and backward linkages, should be incorporated in the estimation model.

Appendix

The correlation coefficients between value-added and explanatory variables are presented at the last column. Basically, for most of the industries Hirschman Herfindahl Index (HHI) has negative correlation to value-added and entropy index has positive correlation to value-added. Both capital stock and employees are strongly correlated to value-added. Means of value-added vary considerably among industries. Industries related to electronics show relatively higher values than traditional industries.

Table A1. Summary statistics of variables

Variables		Observations	Mean	Std. Dev.	Min	Max	Correlation
$R_{j''}$	Highway road length (km)	47	178.9	135.5	18.2	826.2	
P_j	Daytime population	2,138	80,625.5	990,916.5	1,013	4,516,821	
Fish Paste Products (0923)							
	Value-added	1,457	149.9	488.2	0.10	5,493.4	
	Capital stock	1,457	75.3	284.5	0.07	4,208.9	0.791
	Number of employees	1,457	26.3	58.7	1	823	0.792
	Labour productivity	1,457	3.98	5.82	0.05	165.4	
	Capital/labour ratio	1,457	1.38	3.93	0.04	165.4	
	Hirschman Herfindahl index	404	0.69	0.32	0.03	1.00	0.041
	Entropy index	404	0.62	0.75	0.00	3.48	–0.054
	Employees at post code level	1,457	43.32	72.35	1	823	0.735
	Amount of stated capital	1,457	118.44	1,313.09	0.010	23,729.0	0.464
	Value-added/ shipment	1,457	0.52	0.15	0.13	0.91	0.217
Sake: rice wine (1023)							
	Value-added	1,466	177.4	865.7	0.07	17,770.3	
	Capital stock	1,466	124.5	534.6	0.02	9,026.4	0.682
	Number of employees	1,466	17.4	34.5	1	602	0.759
	Labour productivity	1,466	6.58	6.59	0.03	93.6	
	Capital/labour ratio	1,466	3.34	9.30	0.04	235.2	
	Hirschman Herfindahl index	692	0.75	0.28	0.10	1.00	0.010
	Entropy index	692	0.42	0.55	0.00	2.52	–0.013
	Employees at post code level	1,466	17.51	31.04	1	602	0.809
	Amount of stated capital	1,466	21.62	58.15	0.02	1,042.2	0.409
	Value added/ shipment	1,466	0.55	0.15	0.19	0.86	0.251
Fabric mills woven cotton (1141)							
	Value-added	2,014	33.1	193.5	0.02	6,972.5	
	Capital stock	2,014	18.3	195.8	0.02	7,786.2	0.592
	Number of employees	2,014	6.3	14.8	1	256	0.717
	Labour productivity	2,014	3.12	4.17	0.03	71.2	
	Capital/labour ratio	2,014	0.48	2.93	0.03	71.2	
	Hirschman Herfindahl index	219	0.63	0.37	0.01	1.00	–0.381
	Entropy index	219	0.91	1.16	0.00	4.08	0.074
	Employees at postal code level	2,014	23.74	31.65	1	256	0.267
	Amount of stated capital	2,014	32.13	633.23	0.02	22,040.0	0.638
	Value-added/ shipment	2,014	0.64	0.21	0.18	0.94	0.078
Fabric mills woven cotton (1212)							
	Value-added	6,911	26.1	69.09	0.02	1,844.5	
	Capital stock	6,911	6.6	195.8	0.02	7,786.2	0.599
	Number of employees	6,911	10.4	14.8	1	256	0.629
	Labour productivity	6,911	2.28	4.17	0.03	71.2	
	Capital/labour ratio	6,911	0.36	2.93	0.03	71.2	
	Hirschman Herfindahl index	1,046	0.16	0.25	0.01	1.00	0.081
	Entropy index	1,046	0.24	0.15	0.00	3.28	0.242
	Employees at post code level	6,911	19.30	26.16	1	253	–0.367
	Amount of stated capital	6,911	6.15	131.97	0.02	10,040.0	0.129
	Value-added/ shipment	6,911	0.72	0.19	0.08	0.94	–0.107

Table A1. Continued

Variables	Observations	Mean	Std. Dev.	Min	Max	Correlation
Towel products (1296)						
Value-added	459	66.1	181.9	0.12	2,055.9	
Capital stock	459	24.2	86.8	0.09	861.1	0.613
Number of employees	459	9.5	16.9	1	143	0.769
Labour productivity	459	5.01	6.36	0.06	74.2	
Capital/labour ratio	459	0.89	2.94	0.07	35.9	
Hirschman Herfindahl index	66	0.78	0.316	0.016	1.00	–0.364
Entropy index	66	0.49	0.973	0.01	4.87	0.0168
Employees at post code level	459	52.09	75.69	1	268	0.536
Amount of stated capital	459	6.29	99.18	0.01	80.00	0.489
Value-added/ shipment	459	0.53	0.25	0.13	0.89	0.088
Wooden furniture (1411)						
Value-added	7,952	55.1	311.3	0.03	18,899.1	
Capital stock	7,952	18.2	134.4	0.01	4,916.7	0.640
Number of employees	7,952	10.1	26.9	1	663	0.755
Labour productivity	7,952	4.21	3.99	0.03	84.6	
Capital/labour ratio	7,952	0.49	1.95	0.01	72.24	
Hirschman Herfindahl index	1,184	0.59	0.34	0.01	1.00	–0.326
Entropy index	1,184	0.87	0.92	0.00	4.17	0.242
Employees at post code level	1,184	26.85	58.35	1	690	0.456
Amount of stated capital	7,952	90.36	2,832.04	0.01	1,383.4	0.609
Value-added/ shipment	7,952	0.57	0.18	0.12	0.91	0.065
Plastic shoes (2022)						
Value-added	792	39.5	220.4	0.06	4,977.8	
Capital stock	792	6.9	78.3	0.02	1,847.8	0.506
Number of employees	792	9.6	52.8	1	872	0.742
Labour productivity	792	3.32	5.31	0.05	74.8	
Capital/labour ratio	792	0.65	0.82	0.01	13.9	
Hirschman Herfindahl index	90	0.78	0.32	0.05	1.00	–0.318
Entropy index	90	0.53	1.04	0.00	5.47	0.290
Employees at post code level	792	77.26	99.68	1	872	0.487
Amount of stated capital	792	5.23	55.69	0.01	1,300	0.315
Value-added/ shipment	792	0.63	0.25	0.01	0.95	0.225
Clay roofing tile (2231)	792					
Value-added	566	86.5	227.2	0.14	2,773.2	
Capital stock	566	63.8	181.2	0.01	1,566.3	0.739
Number of employees	566	11.9	21.3	1	307.0	0.836
Labour productivity	566	4.49	4.77	0.32	42.7	
Capital/labour ratio	566	1.97	5.10	0.08	41.4	
Hirschman Herfindahl index	139	0.75	0.31	0.01	1.00	–0.156
Entropy index	139	0.50	0.81	0.00	4.57	0.124
Employees at post code level	566	94.95	122.60	1	411	0.245
Amount of stated capital	566	26.82	158.07	0.01	1,817.3	0.425
Value-added/ shipment	566	0.59	0.16	0.02	0.92	0.167
Moulds and dies, parts (2696)						
Value-added	9,943	98.9	372.0	0.06	17,200.1	
Capital stock	9,943	37.2	194.5	0.11	9,211.9	0.701
Number of employees	9,943	12.9	45.6	1	2,455.0	0.789
Labour productivity	9,943	7.03	5.14	0.06	55.8	
Capital/labour ratio	9,943	1.19	3.33	0.04	88.7	
Hirschman Herfindahl index	978	0.56	0.34	0.00	1.00	–0.466
Entropy index	978	1.02	1.04	0.00	5.24	0.430
Employees at post code level	9,943	37.38	59.24	1	2455	0.494
Amount of stated capital	9,943	50.07	927.16	0.01	38,909.1	0.415
Value-added/ shipment	9,943	0.68	0.16	0.02	0.95	0.166

Table A1. Continued

Variables	Observations	Mean	Std. Dev.	Min	Max	Correlation
Bicycles and parts (3091)						
Value-added	369	212.0	1,843.8	0.60	34,654.7	
Capital stock	369	81.1	656.7	0.09	12,099.9	0.705
Number of employees	369	26.0	98.5	1	1259	0.762
Labour productivity	369	4.54	5.39	0.02	38.6	
Capital/labour ratio	369	2.43	3.65	0.07	20.4	
Hirschman Herfindahl index	133	0.848	0.251	0.09	1.00	–0.387
Entropy index	133	0.295	0.568	0.00	2.73	0.521
Employees at post code level	369	31.07	80.23	1	1259	0.848
Amount of stated capital	369	237.24	2,629.9	0.05	35,613.1	0.102
Value-added/ shipment	369	0.65	0.22	0.02	0.85	0.192
Ophthalmic goods, frames (3161)						
Value-added	796	101.8	585.78	0.13	12,544.3	
Capital stock	796	37.9	23.70	0.03	4,932.8	0.753
Number of employees	796	15.2	47.3	1	441	0.846
Labour productivity	796	3.87	4.26	0.08	38.1	
Capital/labour ratio	796	1.62	3.28	0.04	42.6	
Hirschman Herfindahl index	71	0.823	0.295	0.02	1.00	–0.339
Entropy index	71	0.436	0.937	0.00	5.09	0.306
Employees at post code level	796	101.05	131.21	1	494	0.181
Amount of stated capital	796	89.08	1,900.7	0.03	53,204.1	0.556
Value-added/ shipment	796	0.67	0.21	0.02	0.95	0.176
Jewellery products (3211)						
Value-added	1,154	64.5	446.4	0.38	134,963.9	
Capital stock	1,154	22.2	242.9	0.09	7775.6	0.650
Number of employees	1,154	8.7	22.4	1	2550	0.758
Labour productivity	1,154	4.57	6.08	0.05	62.2	
Capital/labour ratio	1,154	2.53	3.55	0.03	92.3	
Hirschman Herfindahl index	254	0.719	0.32	0.02	1.00	–0.481
Entropy index	254	0.598	0.83	0.00	4.27	0.356
Employees at post code level	1,154	19.63	35.33	1	468	0.402
Amount of stated capital	1,154	166.66	3,520.9	0.41	11,5703.0	0.524
Value-added/ shipment	1,154	0.62	0.22	0.15	0.84	–0.197
Communication equipment (2810)						
Value-added	7,338	1,066.4	7,551.7	0.28	146,586.7	
Capital stock	7,338	636.1	4,984.1	0.15	105,108.9	0.771
Number of employees	7,338	83.7	253.1	1	4316	0.797
Labour productivity	7,338	5.87	9.84	0.02	332.4	
Capital/labour ratio	7,338	2.06	5.98	0.01	55.0	
Hirschman Herfindahl index	550	0.76	0.27	0.09	1.00	0.316
Entropy index	550	0.43	0.56	0.00	2.66	–0.027
Employees at post code level	7,338	90.19	267.50	1	4316	0.618
Amount of stated capital	7,338	2,832.8	24,632.5	0.50	33,782.0	0.506
Value-added/ shipment	7,338	0.62	0.22	0.20	0.85	–0.304
Electronic computer (2820)						
Value-added	1,160	1,078.8	5,778.3	0.05	80,960.5	
Capital stock	1,160	300.8	1,597.9	0.01	30,223.2	0.772
Number of employees	1,160	87.9	271.0	1	3672	0.807
Labour productivity	1,160	6.87	14.67	0.09	352.3	
Capital/labour ratio	1,160	2.53	7.45	0.07	226.6	
Hirschman Herfindahl index	467	0.79	0.26	0.09	1.00	0.307
Entropy index	467	0.37	0.52	0.00	2.71	–0.287
Employees at post code level	1,160	73.80	221.39	1	3,672	0.742
Amount of stated capital	1,160	2,083.3	21,094.8	0.20	337,821.1	0.168
Value-added/ shipment	1,160	0.63	0.22	0.21	0.85	0.023

Table A1. Continued

Variables	Observations	Mean	Std. Dev.	Min	Max	Correlation
Electronic parts and devices (2910)						
Value-added	7,338	1,066.4	7,551.7	0.81	278,554.4	
Capital stock	7,338	636.0	4,984.1	0.02	184,360.7	0.805
Number of employees	7,338	83.7	253.1	1	5196	0.808
Labour productivity	7,338	5.87	9.841	0.16	393.6	
Capital/labour ratio	7,338	3.13	5.98	0.06	223.5	
Hirschman Herfindahl index	1,213	0.63	0.31	0.03	1.00	0.275
Entropy index	1,213	0.77	0.76	0.00	4.22	0.351
Employees at post code level	7,338	87.09	275.78	1	5939	0.596
Amount of stated capital	7,338	1,744.2	16,279.6	0.21	32,4625.1	0.592
Value-added/ shipment	7,338	0.63	0.23	0.12	2.51	–0.268

Notes: Value added, capital, labour productivity, and capital/labour ratio are expressed in million yen.
The column of 'Correlation' exhibits simple correlation coefficients between labour productivity and explanatory variables.

References

Aigner DC, Lovell AK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6: 21–37

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

Anderson M, Lööf H (2011) Agglomeration and productivity: Evidence from firm-level data. *Annals of Regional Science* 46: 601–620

Baldwin JR, Beckstead D, Brown WM, Rigby DL (2007) Agglomeration and the geography of localization economies in Canada. *Regional Studies* 42: 117–132

Baldwin JR, Brown WM, Rigby DL (2010) Agglomeration economies: Microdata panel estimation from Canadian manufacturing. *Journal of Regional Science* 50: 915–934

Barrios S, Bertinelli L, Strobl E (2006) Geographic concentration and firm scale: An extension using panel data. *Journal of Regional Science* 46: 733–746

Battese GE, Broca SS (1997) Functional forms of stochastic frontier production functions and models for technical inefficiency effects: A comparison study for wheat farmers in Pakistan. *Journal of Productivity Analysis* 8: 395–414

Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20: 325–332

Battese GE, Corra GS (1977) Estimation of a production frontier model: With application to the pastoral zone of eastern Australia. *Australian Journal of Agricultural Economics* 21: 169–179

Beeson PE, Husted S (1989) Patterns and determinants of productive efficiency in state manufacturing. *Journal of Regional Science* 29: 15–18

Ellison G, Glaeser E (1997) Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy* 105: 889–927

Figueiredo O, Guimaraes P, Woodward D (2009) Localization economies and firm size: Was Marshall right after all? *Journal of Economic Geography* 9: 853–868

Holmes T, Stevens J (2002) Geographic concentration and firm scale. *Review of Economics and Statistics* 84: 682–690

Holmes T, Stevens J (2004) Geographic concentration and firm size: Analysis in an alternative economic geography model. *Journal of Economic Geography* 4: 227–250

Huang CJ, Liu J (1994) Estimation of non-neutral stochastic frontier production function. *Journal of Productivity Analysis* 5: 171–180

Jacobs J (1969) *The economy of cities*. Vintage, New York

Kumbhakar SC, Lovell CAC (2000) *Stochastic frontier analysis*. Cambridge University Press, Cambridge.

Lafourcade M, Mion G (2007) Concentration, agglomeration and the size of plants. *Regional Science and Urban Economics* 37: 46–68

Lall SV, Shalizi Z, Deichmann U (2004) Agglomeration economies and productivity in Indian industry. *Journal of Development Economics* 73: 643–673

Marshall A (1920) *Principles of economics*. MacMillan, London

Ministry of Economy, Trade, and Industry (2007) *Census of manufactures 2005*. Keizai Chosa Kyokai (Economics Research Association), Japan

Ministry of Internal Affairs and Communications (2011) 2009 economic census for business frame, Japan. URL: http://www.stat.go.jp/english/data/e-census/index.htm

Mitra A (1999) Agglomeration economies as manifested in technical efficiency at the firm level. *Journal of Urban Economics* 45: 490–500

Nakamura R (1985) Agglomeration economies in urban manufacturing industries: A case of Japanese cities. *Journal of Urban Economics* 17: 108–124

Otsuka A, Gotoh M, Sueyoshi T (2010) Industrial agglomeration effects in Japan: Productive efficiency, market access, and public fiscal transfer. *Papers in Regional Science* 89: 819–840

Overman HG, Puga D (2009) Labour pooling as a source of agglomeration: An empirical investigation. Working Paper. CEPR, London

Puig-Junoy D (2001) Technical inefficiency and public capital in US states: A stochastic frontier approach. *Journal of Regional Science* 41: 75–96

Rosenthal SS, Strange WC (2001) The determinants of agglomeration. *Journal of Urban Economics* 50: 191–229

Rosenthal SS, Strange WC (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson JV, Thisse J-F (eds) *Handbook of urban and regional economics*, *volume 4*. North Holland, New York

Traistaru I, Iara A (2002) European integration, regional specialization and location of industrial activity in accession countries: Data and measurement. Phare ACE Project P98-1117-R, Center for European Integration Studies, University of Bonn

Tveteras R, Battese GE (2006) Agglomeration externalities, productivity, and technical inefficiency. *Journal of Regional Science* 46: 605–625