Effects of Industrial Agglomeration on Regional Productivity Difference in Japan: an Empirical Study of the New Economic Geography Theory

Kenta Ikeuchi *

August 5, 2013

PRELIMINARY DRAFT; PLEASE DO NOT CITE WITHOUT AUTHOR'S PERMISSION.

Abstract

In this paper, the effects of transportation costs on agglomeration economy and dynamics of industrial location are examined empirically. Combining a spatial demand function derived in the theoretical New Economic Geography (NEG) literature, Krugman (1980), Fujita, Krugman and Venables (1999) and others, with a production function, I propose a revenue production function, which captures the effects of transportation costs on firm's revenue. The suggested revenue production function makes it possible to relate the geographic agglomeration economy with the transportation costs, have not yet done in previous empirical studies. An empirical examination of the model with regional panel data of manufacturing sector in Japan are performed. I estimate the revenue production function including parameters for transportation costs of each industry. The results support the existence of the positive transportation costs and estimated transportation costs for manufacturing products are higher than that for primary sector and lower than that for service sector.

1 Introduction

This paper analyzes spatial effects of industrial geographic locations on regional productivity. According to the theoretical literature in the New Economic Geography (NEG), scale economies and transportation costs create an agglomeration economy (e.g., Fujita et al., 1999; Fujita and Thisse, 2002; Krugman, 1980). The NEG theories have implied that firms located different places face the different demand functions for each other. Klette and Griliches (1996) suggested the inconsistency of scale estimators obtained from production function regressions when firms operate in an incomplete competitive market and prices differ between them. Levinsohn and Melitz (2002) discuss the biases on the productivity estimation in the case of product differentiated market (industry).

As the another aspect of industrial location, a knowledge agglomeration might create an agglomeration economy through the learning from each other firms located nearby which effect used to be

*National Institute of Science and Technology Policy, Japan.

referred to as "knowledge spillovers". Thus, industrial geographic locations affects regional productivity through these two paths; that is the transportation costs and the knowledge spillovers, and the amount of these two effects determines the optimal industrial locations and geographical resource allocation.

Economic efficiency and optimality of industrial locations have been analyzed theoretically (e.g., Baldwin, Forslid, Martin, Ottaviano and Robert-Nicoud, 2003) [1]. There are also many empirical studies on locational effects on regional productivity or growth. Knowledge spillovers effects are investigated mainly in the literature on Industrial Organization (e.g., Alvarez and Molero, 2005; Fosfuri and Ronde, 2004; Henderson, 2007; Monjon and Waelbroeck, 2003; Ornaghi, 2006). The effects of regional economic density on regional productivity are estimated as "agglomeration effects" in the literature on Regional Economics and Urban Economics (e.g., Bode, 2004; Brülhart and Mathys, 2008; Ciccone, 2002; Ciccone and Hall, 1996; Combes, Duranton, Gobillon and Roux, 2008; Tveteras and Battese, 2006). However, there are few empirical studies on the topic which estimate regional productivity model directly derived from the NEG theoretical models [2]. Because of this, the estimation results in previous studies have not been linked straightforward to the theoretical models, thus, it has been unable to evaluate the efficiency and optimality of actual industrial locations.

Unfortunately, the NEG theoretical models are too complicated to estimate straightforward and its nonlinearity caused other computational issues to be solved. This paper challenge these issues. A "tractable" model is derived from the NEG theory. Because the proposed empirical framework directly derived from the NEG theory, it makes possible to evaluate efficiency and optimality of actual industrial locations using estimation results. The proposed framework, then, will be applied to a region-industry level panel data from the Census of Manufacturing in Japan and the spatial effects of transportation costs on plant level productivity will be estimated.

The paper is organized as follows. Section 2 set out an theoretical model from the literatures. In section 3 I discuss about typical data restriction for a researcher and about modification of theoretical model proposed in section 2. Section 4 discusses empirical methodology and the estimation results are reported in Section 5. Section 6 concludes the paper.

2 Theoretical Model

2.1 Production Function

There are assumed to be I industries and R regions. I denote a set of industries by $\mathcal{I} = \{1, 2, ..., I\}$, a set of regions by $\mathcal{R} \equiv \{1, 2, \ldots, R\}$, and a set of N_{ri} firms within region r and industry i by $\mathcal{J}_{ri} \equiv \{1, 2, \ldots, N_{ri}\}$. Production function of firm $j \in \mathcal{J}_{ri}$ is defined by

$$q_j = \Omega_j \ell_j^{\alpha_L^i} k_j^{\alpha_K^i} \prod_{h \in \mathcal{I}} m_{jh}^{\alpha_h^i} \qquad \forall j \in \mathcal{J}_{ri}, \forall r \in \mathcal{R}, \forall i \in \mathcal{I}, \tag{1}$$

[1]Baldwin et al. (2003) used simplified versions of the models of Krugman (1980) and Fujita et al. (1999).

[2]Hanson (2005) and Mion (2004) estimated wage equation of the NEG theory. In addition, Crozet (2004) and Pons, Paluzie, Silvestre and Tirado (2007) estimated labor migration model which was derived from the NEG theory. Davis and Weinstein (2008) analysed home market effects on production location in new economic geography models. See Brakman, Garretsen, Gorter, Horst and Schramm (2009) and Redding (2010) for further literature review.

where q_j represents the quantity produced by firm j, Ω_j is the knowledge (total factor productivity: TFP) of firm j, ℓ_j and k_j are firm j's labor input and capital stock, respectively. m_{jh} is an aggregate of the varieties of individual intermediate inputs of firm j in industry $h \in \mathcal{I}$ defined by a CES function of the form;

$$m_{jh} \equiv \left[\sum_{s \in \mathcal{R}} \sum_{k \in \mathcal{J}_{sh}} \iota_{jk}^{(\sigma_h - 1)/\sigma_h} \right]^{\sigma_h/(\sigma_h - 1)}, \tag{2}$$

where ι_{jk} represents the firm j's intermediate inputs of each available variety produced by firm $k \in \mathcal{J}_{sh}$ (which locates in region s and belonging to industry h). $\sigma_h > 1$ represents the elasticity of substitution between any two intermediate varieties of industry h, *e.g.* ι_{jk} and $\iota_{jk'}$ where $k, k' \in J_h$. α_L^i, α_K^i and α_h^i in (1) and σ_h in (2) are the production technology parameters to be estimated.

2.2 Demand Function

In this subsection, I will set assumptions for consumers' preference and will relate the f.o.b. price of firm j, p_j, to the location of the firm. Also, firm's sales revenue can be relate to the firm's location independently of its production technology and its input level of production factors.

Consumers' demand Following Fujita et al. (1999), consumers' utility function is assumed to be that of the Dixit-Stiglitz model of monopolistic competition (Dixit and Stiglitz, 1977);

$$\mathcal{U}_s = \prod_{i \in \mathcal{I}} Z_{si}^{\mu_i}, \qquad \forall s \in \mathcal{R}, \tag{3}$$

where $\mathcal{U}_s$ represents the utility of consumer living in region s, Z_{si} represents the consumption aggregate of commodity i which is defined by

$$Z_{si} = \left(\sum_{j \in J_i} z_{sj}^{(\sigma_i - 1)/\sigma_i} \right)^{\sigma_i/(\sigma_i - 1)}, \qquad \forall i \in \mathcal{I}, \tag{4}$$

where z_{sj} is the consumption of each variety j and σ_i is the elasticity of substitution between any two varieties of commodity i.

Denoting the consumer's income by Y_s and the price of goods j in region s by p_{js}, consumer in region s maximize utility (3) subject to the budget constraint:

$$\sum_{i \in \mathcal{I}} \sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{J}_{ri}} p_{js} z_{sj} \leq Y_s.$$

As the solution of such utility maximization problem, consumer demand function in region s for the variety $j \in \mathcal{J}_{ri}$ (in industry $i \in \mathcal{I}$ which is produced in region r), therefore, is derived:

$$z_{sj} = \left(\frac{p_{js}}{G_{is}} \right)^{-\sigma_i} \left(\frac{\mu_i Y_s}{G_{is}} \right) \tag{5}$$

where

$$G_{is} \equiv \left[\sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{J}_{ri}} p_{js}^{-(\sigma_i - 1)} \right]^{-1/(\sigma_i - 1)} \tag{6}$$

is the price index of commodity $i \in \mathcal{I}$ in region s, corresponding to the definition of quantity indexes (2) and (4).

Intermediate demand In the same way, solving the profit maximization problem of firm $k \in \mathcal{J}_{sh}$, which locates region s and belongs to industry $h \in \mathcal{I}$, the intermediate demand of firm k for a variety produced by firm $j \in \mathcal{J}_{ri}$, which locates in region r and belongs to industry i can be derived:

$$\iota_{kj} = \left(\frac{p_{js}}{G_{is}} \right)^{-\sigma_i} \left(\frac{\beta_i^h x_k}{G_{is}} \right) \tag{7}$$

where x_k represents the total expenditure of firm k on its intermediate inputs and β_i^h is the share of expenditure on intermediate inputs from industry i, which is assumed to be constant for the firms in an industry.

Total demand Taking the summation of consumer demand function (5) for all regions and the intermediate demand function (7) for all firms in region s, the total demand function for the product of an individual firm j is

$$\begin{aligned} q_{js} &= \left(z_{sj} + \sum_{h \in \mathcal{I}} \sum_{k \in \mathcal{J}_{sh}} \iota_{kj} \right) \\ &= \left(\frac{p_{js}}{G_{is}} \right)^{-\sigma_i} \frac{E_{si}}{G_{is}} \end{aligned} \tag{8}$$

where E_{si} is the total expenditure in region s on the products of industry i which is defined by

$$E_{si} \equiv \mu_i Y_s + \sum_{h \in \mathcal{I}} \sum_{k \in \mathcal{J}_{sh}} \beta_i^h x_k. \tag{9}$$

Thus, equation (8) indicates that the level of demand for a individual firm j in region s depends on the relative price of the firm (relative to the industry level price index), p_{js}/G_{is}, and the industry level total demand in a real term in the region, E_{si}/G_{is}.

2.3 Transportation Costs and Regional Pricing

According to the theoretical NEG models, a scale economy and transportation costs determine industrial geographic locations (Fujita et al., 1999). The scale economy generates a centripetal force for the industrial locations and the transportation costs generates both the centripetal force and dispersion force in the case of the existence of some immobile factors. Therefore, to examine the optimal industrial locations, it is necessary to measure levels of the transportation cost of each commodity. In this subsection, I define a transportation cost function of firms and incorporate it to the production function defined in equation (1).

Transportation cost to deliver the products of firm $j \in \mathcal{J}_{ri}$ which produces a variety of industry i from region r to region s is assumed to be

$$TC_{js} = p_{js} q_{js} \left(\frac{\tau_j(d_{rs}^G) - 1}{\tau_j(d_{rs}^G)} \right) \tag{10}$$

where p_{js} represents the c.i.f. (cost, insurance and freight) price or delivered price of the product of firm j in region s and q_{js} represents the quantity of the product of firm j, which is transported from production region r to consuming region s. The function $\tau_j(d_{rs}^G)$ represents the transportation technology firm j and $d_{rs}^G > 0$ is the (geographic) distance between region r and region s. The transportation technology $\tau(d)$ of (10) is assumed to satisfy the following conditions:

- $\tau(d) > 0$ for any $d > 0$,
- $\tau(0) = 1$,
- $\partial \tau(d)/\partial d > 0$ for any $d > 0$ and
- if $d \to \infty$ then $\tau(d) \to 0$.

In this paper, I specify the distance weight function for transportation technology as follows:

$$\tau_j(d_{rs}^G) = e^{\tau_i d_{rs}^G}. \tag{11}$$

Suppose that firm j determines the c.i.f. price p_{js} to maximize the profit:

$$\sum_{s \in \mathcal{R}} (p_{js} q_{js} - TC_{js}) - C(q_j)$$

where $C(q_j)$ is the production cost of firm j as a function of the firm's total products, $q_j = \sum_{s \in \mathcal{R}} q_{js}$. As the solution for the profit maximization problem under the monopolistic competition assumption, the c.i.f. price of firm j for each region s is given by:

$$p_{js} = p_j \tau_j(d_{rs}^G) = p_j e^{\tau_i d_{rs}^G} \tag{12}$$

where p_j represents the mill or f.o.b. (free on board) price ($p_j \equiv p_{jr}$ where r is the index of the location of firm j). This implies that the c.i.f. price p_{js} for each region s is proportional to the f.o.b. price and τ_j. Also, the price setting rule described by equation (12) implies that the assumption (10) is equivalent to the assumption of an "iceberg" form of transportation costs.

Substituting p_{js} in the total demand function for firm j (8) by (12), demand of region s for firm j is rewritten as:

$$q_{js} = p_j^{-\sigma_i} E_{si} G_{is}^{\sigma_i - 1} e^{-\sigma_i \tau_i d_{rs}^G},$$

so, the total demand for firm j is

$$q_j \equiv \sum_{s \in \mathcal{R}} q_{js} = p_j^{-\sigma_i} \left(\sum_{s \in \mathcal{R}} E_{si} G_{is}^{\sigma_i - 1} e^{-\sigma_i \tau_i d_{rs}^G} \right). \tag{13}$$

Therefore, the inverse demand function is obtained as:

$$p_j = q_j^{-1/\sigma_i} \left(\sum_{s \in \mathcal{R}} E_{si} G_{is}^{\sigma_i - 1} e^{-\sigma_i \tau_i d_{rs}^G} \right)^{1/\sigma_i} \tag{14}$$

which determines the (f.o.b) price of products of firm j in region r in industry i.

2.4 Revenue Production Function

Combining the inverse demand function (14) with the production function (1) yields a "revenue production function":

$$v_j = \phi_{ri}^{1/\sigma_i} \left(\Omega_j \ell_j^{\alpha_L^i} k_j^{\alpha_K^i} x_j^{\alpha_M^i} \prod_{\forall h \in \mathcal{I}} G_{hr}^{-\alpha_h^i} A_i \right)^{(\sigma_i - 1)/\sigma_i} \tag{15}$$

where $v_j \equiv p_j q_j$ represents the sales revenue of firm j and ϕ_{ri} is the demand shifter for industry i in region r defined as,

$$\phi_{ri} \equiv \sum_{s \in \mathcal{R}} E_{si} G_{is}^{\sigma_i - 1} e^{-\sigma_i \tau_i d_{rs}^G}. \tag{16}$$

G_{is} is the price index of commodity i in region r defined in equation (6). Substituting equation (12) for p_{js} in equation (6), G_{is} can be rewritten as:

$$G_{is} = \left[\sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{J}_{ri}} \left(p_j e^{\tau_i d_{rs}^G} \right)^{-(\sigma_i - 1)} \right]^{-1/(\sigma_i - 1)}. \tag{17}$$

Ω_j is the knowledge of firm j, ℓ_j and k_j is the amount of labor inputs and capital stock, respectively. x_j is firm j's expenditure for intermediate input goods h, defined as $x_j \equiv \sum_{h \in \mathcal{I}} G_{hr} m_{jh}$ where G_{hr} is the price index of commodity h in region r and m_{jh} is the quantity of the intermediate inputs produced by industry h used by firm j. Here, I assume that the cost share of intermediate inputs are optimally determined as follows:

$$\frac{G_{hr} m_{jh}}{\sum_{h' \in \mathcal{I}} G_{h'r} m_{jh}} = \frac{\alpha_h^j}{\sum_{h' \in \mathcal{I}} \alpha_{h'}^j}. \tag{18}$$

Then, using equation (18), the quantity of the intermediate inputs m_{jh} in the production function (1) can be substituted by the following equation:

$$\prod_{\forall h \in \mathcal{I}} m_{jh}^{\alpha_h^i} = \left(\frac{x_j^{\alpha_M^j}}{\prod_{\forall h \in \mathcal{I}} G_{hr}^{\alpha_h^j}} \right) \left(\frac{\prod_{\forall h \in \mathcal{I}} (\alpha_h^i)^{\alpha_h^i}}{(\alpha_M^i)^{\alpha_M^i}} \right)$$

where $\alpha_M^i \equiv \sum_{\forall h \in \mathcal{I}} \alpha_h^i$. And A_i in equation (15) is the industry specific constant term, defined as $A_i \equiv \prod_{\forall h \in \mathcal{I}} (\alpha_h^i)^{\alpha_h^i} / (\alpha_M^i)^{\alpha_M^i}$.

2.4.1 Special Cases

To see the characteristics of the derived revenue production function (15), it should be a good examination to look at several extreme special cases.

Case 1 First, let's assume if there is no transportation costs, $\tau_i = 0$ for all $i \in \mathcal{I}$, then demand condition and supply condition (price index) is constant across regions:

$$\phi_{ri} = \sum_{s \in \mathcal{R}} E_{si} \bar{G}_i^{\sigma_i - 1} = \bar{\phi}_i, \ \forall r \in \mathcal{R}$$

and

$$G_{hr} = \left[\sum_{s \in \mathcal{R}} \sum_{j \in \mathcal{J}_{sh}} p_j^{-(\sigma_h - 1)} \right]^{-1/(\sigma_h - 1)} = \bar{G}_h, \ \forall r \in \mathcal{R}.$$

Thus, revenue production function becomes:

$$v_j = \tilde{\Omega}_j \ell_j^{\tilde{\alpha}_L^i} k_j^{\tilde{\alpha}_K^i} x_j^{\tilde{\alpha}_M^i} \tilde{A}_i \tag{19}$$

where $\tilde{A}_i = \bar{\phi}_i^{1/\sigma_i} \left[A_i \prod_{\forall h \in \mathcal{I}} \bar{G}_h^{\alpha_h^i} \right]^{(\sigma_i - 1)/\sigma_i}$, $\tilde{\Omega}_j = \Omega_j (\sigma_i - 1)/\sigma_i$, $\tilde{\alpha}_L^i = \alpha_L^i (\sigma_i - 1)/\sigma_i$, $\tilde{\alpha}_K^i = \alpha_K^i (\sigma_i - 1)/\sigma_i$ and $\tilde{\alpha}_M^i = \alpha_M^i (\sigma_i - 1)/\sigma_i$. In this case, the revenue production function is almost the same as production function (1). Thus, it is difficult to identify the parameter for elasticity of substitution, σ, by the production function estimation under the condition of extremely low transportation costs.

Case 2 Next case is where the products are perfectly homogeneous and price elasticity of demand is infinite, $\sigma = \infty$ for all $i \in \mathcal{I}$, then $G_{ir} = 1$ and $\phi_{ri} = E_{ri}$ for all $r \in \mathcal{R}$. Thus,

$$v_j = \Omega_j \ell_j^{\alpha_L^i} k_j^{\alpha_K^i} x_j^{\alpha_M^i} A_i. \tag{20}$$

Also in this case, as similar to Case 1, revenue production function is the perfectly same as production function (1). Thus, we cannot identify transportation cost parameter τ from the production function estimation if the products are perfectly homogeneous and price elasticity of demand $\sigma = \infty$. Because the revenue production function in Case 2 has the same form of that function in Case 1, unless other information is not available, also we can not distinguish Case 2 from Case 1. In other words, both in these two cases firms revenue do not depend on demand or supply agglomeration. Thus, in my empirical analysis I treat these cases as a null model.

Case 3 The opposite assumption to Case 1 is infinite transportation cost, $\tau_i = \infty$ for all $i \in \mathcal{I}$. Then, $\phi_{ri} = E_{ri} G_{ir}^{\sigma_i - 1}$ and $G_{ir}^{-(\sigma_i - 1)} = \sum_{j \in \mathcal{J}_{ri}} p_j^{-(\sigma_i - 1)}$ for all $s \in \mathcal{R}$. Thus,

$$v_j = \left(\frac{E_{ri} G_{ir}^{(\sigma_i - 1)}}{\prod_{\forall h \in \mathcal{I}} G_{hr}^{\alpha_h^i (\sigma_i - 1)}} \right)^{1/\sigma_i} \left(\Omega_j \ell_j^{\alpha_L^i} k_j^{\alpha_K^i} x_j^{\alpha_M^i} A_i \right)^{(\sigma_i - 1)/\sigma_i}.$$

In this case, although regional expenditure and price index, $E_{r.}$ and $G_{r.}$, only affect on the revenue of firms within the region itself ($r(j) = r$), we recover σ and α's by an estimation of the equation.

Case 4 Lastly, as the opposite case to Case 2, let's think about a case in which each firm has a perfect monopoly power, in other words, the price elasticity of demand $\sigma = 1$ for all $i \in \mathcal{I}$. Then, $G_{ir} = 1$ and $\phi_{ri} = \sum_{s \in \mathcal{R}} e^{-\tau_i d^G_{rs}} E_{si}$ for all $r \in \mathcal{R}$. Thus,

$$v_j = \phi_{ri} = \sum_{s \in \mathcal{R}} e^{-\tau_i d^G_{rs}} E_{si}.$$

In this case, price index, G, is not varied across regions but demand for the firms, ϕ, does. Moreover, revenue of the firms are not depended on their factor inputs but only depended on demand factor, ϕ, which is affected by difference of expenditure, E, across region. It indicates that it is difficult to recover the output elasticity to factor inputs from revenue production function estimation when the demands for products are perfectly sensitive to its price.

3 Data

3.1 Data Sources

I obtain the data from several statistics in Japan for year 1996-2006. First, information on the production inputs and outputs of the firms in manufacturing industries at regional level are obtained from the "Census of Manufacturers", which is conducted by the Ministry of Economy, Trade and Industry in Japan. This data contains the information on the all plants located in Japan at least with 4 employees. Data at the 2 digit Japanese Standard Industrial Classification (JSIC) level by city level (in Japanese "shi", "ku", "cho" and "son" level) data is available. 2 digit JSIC includes 22 industries. The number of cities in Japan is almost 2000 in the latest year of the observation years[3].

Second, for the estimates of the input coefficients of each industry, I use the "Input-Output Tables for Japan" constructed by the Ministry of Internal Affairs and Communications in Japan. Because this tables are constructed for each 5 years, I interpolate the input coefficients in the intermediate years.

Third, for regional distribution of the workers of non-manufacturing industries, I use the "Establishment and Enterprise Census" conducted by the Ministry of Internal Affairs and Communications in Japan. This census updated for each 3 or 5 years and contains the information on all establishments (excluded for self employments in the primary sector; agriculture, forestry and fishery). Also, I interpolate the share of the number of workers employed in each industry in the region in the intermediate years.

3.2 Further Assumptions for Data Restrictions

In this paper, I have several data restrictions (which might be common for other researchers);

[3]From the mid of 1990s, Japanese administrative division of the regions has been restructured and hundreds of regions were merged each other during this period. To assure the consistency I use the latest (and the largest-meshed) classification for whole period.

1. Individual firm-level (micro level) data can not be obtained but only an aggregate (region-industry level) data can be obtained.

2. Price index in the regional level for each industry cannot be obtained.

3. Regional input and output data can be obtained only for manufacturing industries but can not be obtained for non-manufacturing industries.

3.3 Homogeneity of Firms

In order to estimate the models discussed in the previous section from an aggregated region-industry level dataset instead of firm-level micro data, I have to set the following assumptions.

- Production technologies of the firms are the same in each industry.

$$\alpha_L^j = \alpha_L^i,\ \alpha_K^j = \alpha_K^i,\ \alpha_h^j = \alpha_h^i,\ \forall h \in \mathcal{I}, \forall j \in \mathcal{J}_{ri}, r \in \mathcal{R}. \tag{21}$$

- Production input quantity for each factor is the same across the firms located in the same region and belonging to the same industry.

$$\ell_j = \ell_k, k_j = k_k, m_{jh} = m_{kh}, \forall h \in \mathcal{I}, \forall j, k \in \mathcal{J}_{ri}. \tag{22}$$

- Unobserved efficiency is the same across firms located in the same region and belonging to the same industry.

$$\Omega_j = \Omega_{ri}, \forall j \in \mathcal{J}_{ri}. \tag{23}$$

Then, the region-industry level aggregation of the model (1) is

$$q_j = \frac{Q_{ri}}{N_{ri}} = \Omega_{ri} L_{ri}^{\alpha_L^i} K_{ri}^{\alpha_K^i} \prod_{\forall h \in \mathcal{I}} M_{rih}^{\alpha_h^i}\ \forall j \in \mathcal{J}_{ri} \tag{24}$$

where Q_{ri}, L_{ri}, K_{ri} and M_{ri} represents the quantities of the total output and production inputs of the firms in industry i in region r. From equations (15) and (24), we obtain the region-industry aggregated revenue production function,

$$V_{ri} = A_i^{(\sigma_i-1)/\sigma_i} \frac{\phi_{ri}^{1/\sigma_i}}{\prod_{\forall h \in \mathcal{I}} G_{hr}^{\alpha_h^i(\sigma_i-1)/\sigma_i}} \left(\Omega_{ri} L_{ri}^{\alpha_L^i} K_{ri}^{\alpha_K^i} X_{ri}^{\alpha_M^i}\right)^{(\sigma_i-1)/\sigma_i} \tag{25}$$

3.4 Regional Price Index

Although the price index of industry h in region r, G_{hr}, is defined by equation (17), that data is rarely obtained. Therefore, using the inverse demand function (14), I can rewrite the f.o.b price of firm j, p_j, as:

$$p_j = \left(\frac{\phi_{ri}}{v_j}\right)^{1/(\sigma_i-1)} = \left(\frac{\phi_{ri}}{V_{ri}/N_{r'i}}\right)^{1/(\sigma_i-1)} \tag{26}$$

Because $p_j = v_j/q_j$ and $q_j = p_j^{-\sigma_i}\phi_{ri}$, $p_j = v_j p_j^{\sigma_i}\phi_{ri}^{-1}$.

Then, inserting this equation into (17), we obtain the price index of commodity i in region s:

$$
\begin{aligned}
G_{is} &= \left(\sum_{\forall r' \in \mathcal{R}} \sum_{\forall k \in \mathcal{J}_{ri}} \frac{N_{r'i} v_k}{\phi_{r'i}} e^{-(\sigma_i - 1)\tau_i d_{sr'}}\right)^{-1/(\sigma_i - 1)} \\
&= \left(\sum_{\forall r' \in \mathcal{R}} \frac{V_{r'i}}{\phi_{r'i} e^{(\sigma_i - 1)\tau_i d_{sr'}}}\right)^{-1/(\sigma_i - 1)}. \qquad (27)
\end{aligned}
$$

The equation suggests that the regional price index, G_{is}, for a commodity i is higher in region s if there is larger supply for the commodity, $V_{i.}$, and fewer demand for the commodity, $\phi_{i.}$, in the region itself, s, or its neighboring regions, r' with small $d_{sr'}$. This indicates that lack of the geographic competition leads to a higher price in the region.

3.5 Non-manufacturing Industries

Typically, detailed data of production inputs and output for non-manufacturing sector cannot be obtained at the regional level. Even if we ignore to estimate the production function of non-manufacturing firms, in order to estimate the regional revenue production function for manufacturing sector, we need regional demand potential ϕ and the regional factor price index G for the manufacturing industries, we need regional revenue and expenditure for intermediate inputs of non-manufacturing sector.

Suppose that at the national level revenue, V_i, and total intermediate expenditure, X_i, of non-manufacturing industry i are observable and assuming the their regional level amounts are proportionate to the number of workers employed in the region in the industry. Specifically, I approximate the revenue of the firms in region r in industry i by

$$
V_{ri} = \frac{L_{ri}}{\sum_{\forall r \in \mathcal{R}} L_{ri}} V_i, \quad \forall i \in \mathcal{I}_S \qquad (28)
$$

and also the expenditure for intermediate inputs of the firms in region r in industry i by

$$
X_{ri} = \frac{L_{ri}}{\sum_{\forall r \in \mathcal{R}} L_{ri}} X_i, \quad \forall i \in \mathcal{I}_S \qquad (29)
$$

where $\mathcal{I}_S$ is the set of non-manufacturing industries. In most cases, researchers can observe the number of workers at the detailed regional level even for non-manufacturing industries.

4 Estimation Method

4.1 Revenue Production Function Estimation

Taking natural logarithms of the both sides of equation (25) and adding a time dimension t, the following equation is obtained for all regions $r \in \mathcal{R}$ and all manufacturing industries $i \in \mathcal{I}_M$,

$$\ln V_{ri}^{(t)} = \tilde{\alpha}_0^i + \tilde{\alpha}_L^i \ln L_{ri}^{(t)} + \tilde{\alpha}_K^i \ln K_{ri}^{(t)} + \tilde{\alpha}_M^i \ln X_{ri}^{(t)} + \frac{\ln \phi_{ri}^{(t)}}{\sigma_i} - \tilde{\alpha}_M^i \sum_{\forall h \in \mathcal{I}} \hat{\beta}_h^i \ln G_{hr}^{(t)} + \omega_{ri}^{(t)} \tag{30}$$

$t = 1, 2, ..., T$, where

$$\tilde{\alpha}_0^i \equiv \frac{\sigma_i - 1}{\sigma_i} \ln A_i,\ \tilde{\alpha}_L^i \equiv \frac{\sigma_i - 1}{\sigma_i} \alpha_L^i,\ \tilde{\alpha}_K^i \equiv \frac{\sigma_i - 1}{\sigma_i} \alpha_K^i,\ \tilde{\alpha}_M^i \equiv \frac{\sigma_i - 1}{\sigma_i} \sum_{\forall h \in \mathcal{I}} \alpha_h^i,\ \omega_{ri}^{(t)} \equiv \frac{\sigma_i - 1}{\sigma_i} \ln \Omega_{ri}^{(t)}.$$

ϕ_{ri} is the demand shifter for the firms in region r in industry i defined in equations (16):

$$\phi_{ri}^{(t)} = \sum_{\forall s \in \mathcal{R}} \frac{E_{si}^{(t)}}{e^{\sigma_i \tau_i d_{rs}}} \left(G_{is}^{(t)} \right)^{\sigma_i - 1} \tag{31}$$

where E_{si} is expenditure in region s for commodity i defined as:

$$E_{si}^{(t)} = \sum_{h \in \mathcal{I}} \left[\left(\frac{\hat{\mu}_i \tilde{\alpha}_L^h + \hat{\beta}_i^h \tilde{\alpha}_M^h}{\gamma_h} \right) V_{sh}^{(t)} \right]. \tag{32}$$

The price index of commodity h in region r is given by:

$$G_{hr}^{(t)} = \left(\sum_{\forall r' \in \mathcal{R}} \frac{V_{r'h}^{(t)}}{\phi_{r'h}^{(t)} e^{(\sigma_h - 1)\tau_h d_{rr'}}} \right)^{-1/(\sigma_h - 1)}. \tag{33}$$

Econometric Issues Consistency of the estimator for the parameters rely on assumption on the conditional mean of the unobserved efficiency term, $\omega_{ri}^{(t)}$. According to Wooldridge (2002), in the non-linear regression if,

$$E\left(\omega_{ri}^{(t)} \middle| \frac{\partial \ln V_{ri}^{(t)}}{\partial \tau_h}, \frac{\partial \ln V_{ri}^{(t)}}{\partial \sigma_h}, \ln L_{ri}^{(t)}, \ln K_{ri}^{(t)}, \ln M_{ri}^{(t)} \right) = 0, \tag{34}$$

then the non-linear least squares (NLS) estimator is consistent, where the NLS estimator minimizes the objective function:

$$O_{NLS}(\boldsymbol{\theta}, \tilde{\boldsymbol{\alpha}}) = \sum_r \sum_i \sum_t \left(\ln V_{ri}^{(t)} - h_{ri}^{(t)}(\boldsymbol{\theta}) - \mathbf{X}_{ri}^{(t)} \tilde{\boldsymbol{\alpha}}_i \right)^2 \tag{35}$$

where

$$\boldsymbol{\theta} = (\tau_1\ \tau_2\ \cdots\ \tau_I\ \sigma_1\ \sigma_2\ \cdots\ \sigma_I)',$$

$$\tilde{\boldsymbol{\alpha}}_i \equiv \left(\tilde{\alpha}_L^i\ \tilde{\alpha}_K^i\ \tilde{\alpha}_M^i \right)',$$

$$h_{ri}^{(t)}(\boldsymbol{\theta}) \equiv (1/\sigma_i)\ln\phi_{ri}^{(t)}(\tau_i,\sigma_i) - \tilde{\alpha}_M^i \sum_{\forall h\, in \mathcal{I}} \hat{\beta}_h^i \ln G_{hr}^{(t)}(\tau_h,\sigma_h),$$

and

$$\mathbf{X}_{ri}^{(t)} \equiv (\ln L_{ri}\ \ \ln K_{ri}\ \ \ln X_{ri})\,.$$

Unfortunately, as the large body of the econometric literature on production function estimation has been suggested, the condition in which equation (34) hold cannot be met in the practice. To obtain consistent estimators when the condition equation (34) is violated, I decompose the term of unobserved efficiency of the firms in region r in industry i in year t into three components as follows;

$$\omega_{ri}^{(t)} = \mu_i^{(t)} + \bar{\omega}_{ri} + u_{ri}^{(t)} \tag{36}$$

where $\mu_i^{(t)}$ is the industry-year specific efficiency shock, $\bar{\omega}_{ri}$ is persistent efficiency difference across region in each industry and $u_{ri}^{(t)}$ is a time dependent region specific efficiency shock for each industry. By including industry-year dummies as explanatory variables, the first component can be controlled out jointly with α_0^i.

Next, if,

$$E\left(\bar{\omega}_{ri}\left|\frac{\partial \ln V_{ri}}{\partial \boldsymbol{\theta}}, \mathbf{X}_{ri}^{(t)}\right.\right) \neq 0, \tag{37}$$

then the NLS estimator is no longer consistent. In order to conduct consistent estimation, at least we have to use the fixed effects (FE) or the first difference (FD) estimator in this case. The FE estimator minimizes the objective function:

$$O_{FE}(\boldsymbol{\theta},\tilde{\boldsymbol{\alpha}}) = \sum_r\sum_i\sum_t \left(\ln \ddot{V}_{ri}^{(t)} - \ddot{h}_{ri}^{(t)}(\boldsymbol{\theta}) - \ddot{\mathbf{X}}_{ri}^{(t)}\tilde{\boldsymbol{\alpha}}_i\right)^2 \tag{38}$$

where the variable with "¨" is demeaned, e.g. $\ddot{x} = x_{ri}^{(t)} - \overline{x_{ri}}$. The FD estimator minimizes the least squares of the error term using the first difference of original dataset:

$$O_{FD}(\boldsymbol{\theta},\tilde{\boldsymbol{\alpha}}) = \sum_r\sum_i\sum_t \left(\Delta\ln V_{ri}^{(t)} - \Delta h_{ri}^{(t)}(\boldsymbol{\theta}) - \Delta\mathbf{X}_{ri}^{(t)}\tilde{\boldsymbol{\alpha}}_i\right)^2 \tag{39}$$

If,

$$E\left(u_{ri}^{(t)}\left|\frac{\partial \ln V_{ri}}{\partial \boldsymbol{\theta}}, \mathbf{X}_{ri}^{(t)}\right.\right) \neq 0 \tag{40}$$

then the FE and FD estimators are no longer consistent as well as the NLS estimator. In addition to equation (37), in order to conduct consistent estimation under the condition (40), we have to use some instrumental variable estimator via GMM estimation technique. The GMM estimator minimizes the objective function:

$$O_{GMM}(\boldsymbol{\theta},\tilde{\boldsymbol{\alpha}}) = \left(\sum_r\sum_i \mathbf{u}_{ri}'\mathbf{W}_{ri}\right)\mathbf{A}\left(\sum_r\sum_i \mathbf{u}_{ri}'\mathbf{W}_{ri}\right)' \tag{41}$$

where $\mathbf{u}_{ri}$ is a $(\tilde{T}\times 1)$ vector of residuals for region r industry i, $\mathbf{W}_{ri}$ is a $(\tilde{T}\times P)$ matrix of instrumental variables for region r industry i and $\mathbf{A}$ is a $P \times P$ positive definite matrix called weighting matrix. For the specification of the residual vector $\mathbf{u}$, effective choice of the instruments $\mathbf{W}$ and an efficient estimation of weighting matrix $\mathbf{A}$, I use the difference GMM (DIF-GMM) developed by Arellano and Bond (1991) and the system GMM (SYS-GMM) developed by Blundell and Bond (1998). In summary, the DIF-GMM method use the first differenced residuals $\Delta u_{ri}^{(t)}$ for the element of $\mathbf{u}_{ri}$ and lagged instrumental variables in level $\mathbf{Z}_{ri}^{(t-s)}$, $\mathbf{Z}_{ri}^{(t-s-1)}$, ..., $\mathbf{Z}_{ri}^{(1)}$ for $\mathbf{W}_{ri}^{(t)}$. In addition, the SYS-GMM method add the residuals in level $u_{ri}^{(t)}$ into $\mathbf{u}_{ri}$ and the lagged difference of instrumental variables $\Delta\mathbf{Z}_{ri}^{(t-s)}$ into $\mathbf{W}_{ri}^{(t)}$.

Estimation Procedure The parameters of the model can be estimated by the following estimation procedure. Since the both equation (31) and (33) are non-linear functions and they are dependent on the unknown variable ϕ and G each other, the Gauss-Newton method is employed and a fixed point iteration is nested.

1. Initialize the parameters for transportation cost and elasticity of substitution, $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{\sigma}}$, respectively.
2. Loop the following steps until objective function, $O\left(\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\sigma}}, \hat{\tilde{\boldsymbol{\alpha}}}\right)$, is minimized.
   (a) Initialize $\hat{\phi}_{ri}^{(t)}$ for each r, i and t and loop the following steps until $\hat{\phi}_{ri}^{(t)}$ and $\hat{G}_{ri}^{(t)}$ do not changed for all r, i and t.
      - Given $\hat{\phi}_{ri}^{(t)}$ for all r and i, calculate $\hat{G}_{ri}^{(t)}$ for each r, i and t using equation (33).
      - Given $\hat{G}_{ri}^{(t)}$ for all r and i, update $\hat{\phi}_{ri}^{(t)}$ for each r, i and t using equation (31).

   (b) Using $\hat{\phi}$, $\hat{G}$ and $\hat{\sigma}$, estimate the parameters of linear-part of the revenue production function (30), *i.e.* $\tilde{\alpha}_L^i$, $\tilde{\alpha}_K^i$ and $\tilde{\alpha}_M^i$, $\forall i \in \mathcal{I}_M$, to minimize the augmented objective function,
   $$\tilde{O}\left(\tilde{\boldsymbol{\alpha}}|\hat{\phi}, \hat{G}, \hat{\sigma}\right). \tag{42}$$
   (c) Calculate the value of objective function $\hat{O} = O\left(\hat{\boldsymbol{\theta}}, \hat{\tilde{\boldsymbol{\alpha}}}\right)$.

   (d) Calculate the derivatives of the objective function $\hat{\mathbf{J}} = \begin{pmatrix} \frac{\partial O\left(\hat{\boldsymbol{\theta}},\hat{\tilde{\boldsymbol{\alpha}}}\right)}{\partial \boldsymbol{\theta}'} & \frac{\partial O\left(\hat{\boldsymbol{\theta}},\hat{\tilde{\boldsymbol{\alpha}}}\right)}{\partial \tilde{\boldsymbol{\alpha}}'} \end{pmatrix}$.

   (e) Update the parameters $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{\sigma}}$ using the objective value $\hat{O}$ and the Jacobian $\hat{\mathbf{J}}$.

5 Results

5.1 Data Description

I construct a panel of Japanese regional data from Census of Manufactures, Establishment and Enterprise Census, Population Census and Input-Output Table. This dataset was composed of 1,928 Japanese regions and 22 manufacturing industries (two digits) and 16 non-manufacturing sectors (one digits) spanning the period 1996–2006. Summary statistics are provided in Table 2–6.

5.2 Results of Production Function Estimation

Table 1 reports the estimation results of the revenue production function. This table only reports the estimated values of the parameters in the non-linear part of the model, τ and σ. Although, the elasticity of outputs with respect to factor labor, capital and intermediate inputs are jointly estimated for each 2 digit industry, their estimation results do not appear in the table to avoid an excessive complexity. The first column is the results of the non-linear least squares (NLS) estimation, the second column is the results of the fixed effects (FE) model, and the third column is the system GMM (SGMM) results. The parameters for transportation costs τ are significantly positive for the manufacturing sector in all methods. Although the transportation costs parameters for the primary and service sector are negative in the NLS and FE model, respectively, in the SGMM estimation results parameter for the service sector becomes significantly positive and its magnitude is higher than that for manufacturing sector. Figure 1 illustrates the estimated transportation cost function for each sector.

Elasticity of substitution σ are also estimated for each sector. In the SGMM results, they ranges from 16 for manufacturing to 58 for service sector, which indicates that degree of differentiation of the products in manufacturing sector is higher than primary and service sector. The results of GMM distance statistic test indicates we could not reject a null hypothesis, H0: $\tau = 0$ or $\sigma = \infty$ for all industries, while it is rejected in the NLS and FE model.

Table 1: Estimation results of the revenue production function

		NLS	FE	SGMM
Primary sector	τ	-0.0007	-0.0247	0.0065^{***}
		[0.0101]	[0.0306]	[0.0022]
	σ	33.3211^{***}	29.9847^{***}	31.0388
		[0.0001]	[0.0777]	[25.4835]
Manufacturing	τ	0.0116^{***}	0.0166^{***}	0.0155^{***}
		[0.0009]	[0.006]	[0.0003]
	σ	34.0455^{***}	30.0461^{***}	16.1425^{***}
		[2.2607]	[0.2249]	[0.0793]
Service	τ	0.0406^{***}	-0.0354^{*}	0.0202^{***}
		[0.0038]	[0.0204]	[0.0019]
	σ	11.5428^{***}	30.0095^{***}	58.0532^{***}
		[2.4326]	[9.0884]	[16.1928]
H0: $\tau = 0$ or $\sigma = \infty$		$F(6, 116197) = 8920.6^{***}$	$F(6, 102782) = 2.6^{**}$	$\chi^2(6) = 1.2$

Robust standard errors for NLS and FE and ordinary standard errors for SGMM are in the brackets. ***, ** and * significant at $p < 0.01$, $p < 0.05$ and $p < 0.1$, respectively. In the SGMM estimation, 6 years and more lagged lnL, lnK and lnM are used as "GMM-type" instruments and derivatives with respect to τ and σ are treated as standard instruments without any lags. Sargan test statistic for SGMM is 2006.691 ($p = 0.996$) and the p-value of the Arellano-Bond test for AR(4) in first differences is 0.0292.

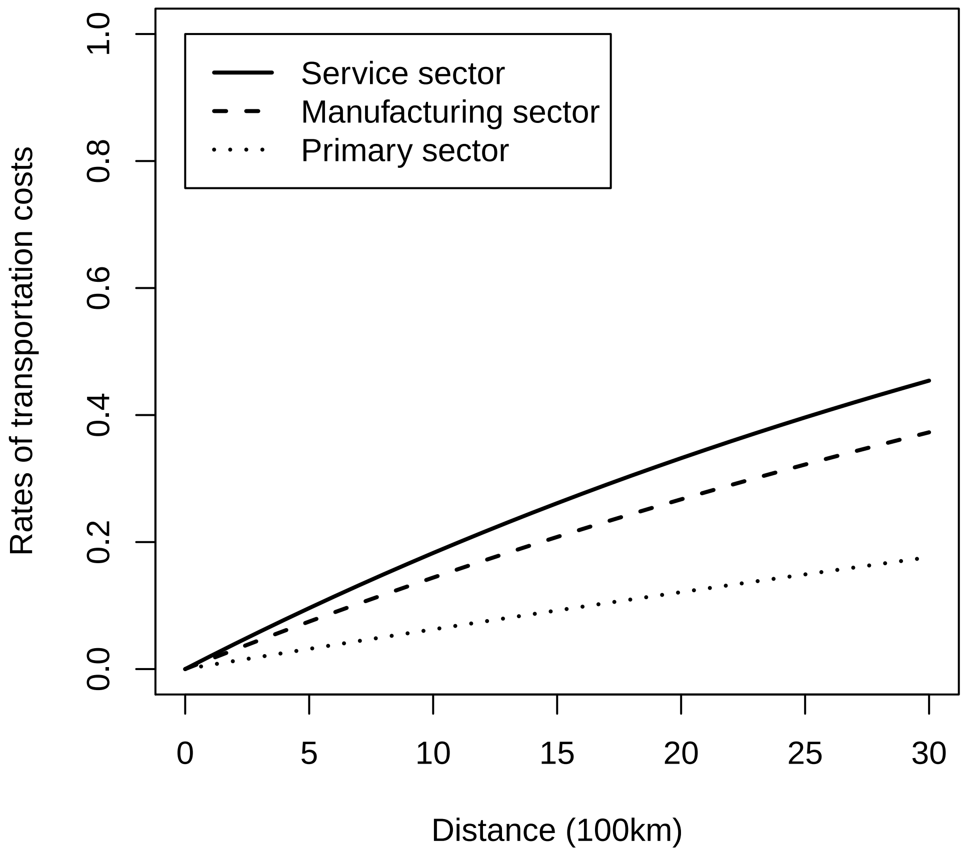

Figure 1: Estimated transportation function

6 Conclusion

In this paper, the effect of transportation costs on agglomeration economy were examined empirically. Combining a spatial demand function derived in the theoretical New Economic Geography (NEG) literature, Krugman (1980), Fujita et al. (1999) and others, with a production function, I proposed a revenue production function, which captures the effects of transportation costs on firm's revenue. Since those spatial effects are generated by transportation costs on firms' own product and its intermediate goods, the suggested revenue production function makes it possible to relate the geographic agglomeration economy with the transportation costs, something not done in previous empirical studies.

I performed an empirical examination of the model with regional panel data of manufacturing sector in Japan. The results of revenue production function estimation show significantly and robust positive transportation costs for the manufacturing products. In addition, a consistent GMM estimation results show an evidence of positive transportation costs not only for the outputs of the manufacturing sector but also for the outputs of the primary and service sector. These results indicate that the efficiency of the manufacturing firms depends on the access to the markets and the access to the intermediate goods supply.

However, there are several remaining issues for future research. First, in order to investigate the effective regional and location policy, empirical analyses of the dynamics of the location of firms and labors are additionally needed. From the results of the paper we can only perform comparative statics. In the dynamic perspective we have to take it account for the other important aspects, e.g. relocation cost, entry cost or time lags. They can be analysed only in dynamic models of location choice. Second, this paper ignores the export and import activities and the rolls of trade hubs, e.g. harbors, airports or train stations. Since the distance from such trade hubs should affects the transportation costs as well as the efficiency of firms, it is necessary to control for the effects of the distance from trade hubs. Third, this paper ignores the knowledge spillovers effects or spillover effects of research and development investments on productivity. Knowledge spillovers effects might be correlated with the market access and supply access which are examined in this paper. Thus, it should be necessary to controlled for the effects of knowledge agglomeration.

Table 2: no. of regions

Industry	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
F09	752	750	753	750	752	754	756	763	798	847	876
F10	356	344	366	360	365	346	321	334	336	394	410
F11	321	312	315	301	301	292	276	278	272	294	307
F12	682	668	677	664	662	643	623	628	635	708	691
F13	583	569	570	555	551	521	497	508	498	562	567
F14	612	601	600	586	580	574	540	548	529	595	580
F15	500	496	505	492	488	484	472	470	481	521	520
F16	681	679	693	685	688	682	669	673	676	721	717
F17	378	377	390	384	384	381	376	376	386	407	422
F18	92	90	98	89	84	84	72	76	69	77	80
F19	585	587	598	589	590	590	580	588	609	664	677
F20	251	238	250	248	242	245	232	236	223	257	264
F21	172	162	169	158	144	137	119	117	96	117	112
F22	672	663	679	667	660	657	647	647	672	735	757
F23	342	334	349	332	330	315	302	315	304	354	366
F24	282	276	284	268	263	264	255	254	249	286	295
F25	717	715	716	713	714	711	706	719	737	799	815
F26	672	678	679	676	679	673	665	679	701	764	781
F27	665	661	665	662	664	655	600	601	625	682	693
F30	498	491	494	490	491	493	474	488	508	561	586
F31	336	334	348	332	325	328	310	306	314	345	350
F32	500	516	548	533	529	520	532	542	565	621	564

Table 3: total number of establishments (1 thousand)

Industry	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
F09	27.6	26.7	28.4	26.8	26.5	25.3	24.1	24.6	24.1	27.5	27.5
F10	2.6	2.5	2.7	2.5	2.6	2.4	2.3	2.4	2.4	3.1	3.2
F11	10.1	9.5	9.1	8.2	7.6	6.9	6.1	6.1	6.0	6.7	6.4
F12	20.1	18.8	19.4	17.1	15.8	13.7	11.9	12.0	10.8	12.2	11.3
F13	8.4	7.9	7.8	7.2	6.9	6.2	5.6	5.7	5.5	6.8	6.5
F14	11.0	10.6	10.4	9.6	9.2	8.6	7.5	7.7	7.0	8.0	7.3
F15	8.2	7.8	8.3	7.7	7.6	7.1	6.6	6.7	6.5	7.0	6.8
F16	23.3	22.7	24.8	22.8	22.4	20.8	17.3	17.6	16.3	16.7	15.6
F17	3.7	3.7	3.9	3.7	3.7	3.6	3.5	3.5	3.6	3.8	3.9
F18	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.3
F19	13.5	13.2	14.2	13.3	13.4	12.7	12.1	12.7	12.5	14.2	13.9
F20	3.3	3.2	3.3	3.1	3.0	2.8	2.5	2.6	2.4	2.7	2.6
F21	3.4	3.2	3.1	2.9	2.6	2.5	2.1	2.1	1.8	1.9	1.7
F22	10.9	10.5	10.8	10.2	9.9	9.4	8.8	8.8	8.9	10.5	10.6
F23	4.2	4.1	4.2	3.8	3.8	3.6	3.4	3.5	3.3	3.7	3.7
F24	2.4	2.4	2.5	2.3	2.3	2.1	2.0	2.1	1.9	2.2	2.3
F25	36.1	35.1	36.6	33.3	33.5	31.3	29.0	30.3	28.7	32.1	30.8
F26	32.3	31.9	33.7	30.7	31.5	29.1	27.1	28.5	27.6	31.3	30.8
F27	21.0	20.4	21.1	19.5	19.4	17.4	15.7	15.9	15.4	17.3	17.0
F30	9.8	9.5	10.0	9.4	9.3	8.9	8.6	9.1	8.9	10.4	10.4
F31	4.2	4.1	4.4	4.0	4.0	3.7	3.4	3.5	3.3	3.6	3.4
F32	7.8	8.2	9.3	8.2	8.4	7.3	7.1	7.6	8.0	8.9	6.8

Table 4: regional no. of employees per establishment

Industry	n.obs	mean	sd	min	median	max
F09	8551	34.424	28.160	5.214	26.625	429.333
F10	3932	25.754	23.259	4.000	17.745	242.250
F11	3269	21.550	22.821	2.600	14.333	246.333
F12	7281	17.888	13.570	3.800	13.833	307.667
F13	5981	14.117	12.287	4.000	11.000	302.000
F14	6345	14.169	24.525	2.400	9.167	767.857
F15	5429	29.041	30.687	4.333	21.143	426.750
F16	7564	19.144	14.534	4.000	15.643	242.800
F17	4261	68.417	79.459	4.000	48.273	1597.000
F18	911	33.067	57.313	4.667	12.000	423.000
F19	6657	26.883	20.417	2.500	21.111	291.250
F20	2686	42.002	61.915	4.333	19.325	501.750
F21	1503	13.745	12.002	4.000	10.615	175.000
F22	7456	23.593	27.694	3.600	17.098	838.000
F23	3643	47.800	69.393	4.333	25.565	920.857
F24	2976	47.282	69.823	4.000	26.000	1227.333
F25	8062	18.799	14.001	3.800	15.500	294.333
F26	7647	31.881	34.257	3.500	23.167	568.667
F27	7173	68.105	85.247	1.857	46.286	1693.800
F30	5574	62.039	118.658	4.000	30.398	2324.667
F31	3628	35.625	45.950	4.333	21.667	982.333
F32	5970	17.566	27.289	0.644	10.944	632.800

Table 5: consumption expenditure for each industry (1 trillion yen)

Industry	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
A	5.4	5.5	5.3	5.3	5.3	5.3	5.3	4.7	4.7	4.8	4.8
B	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.2	0.3	0.3
C	1.2	1.2	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
D	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
E	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F09	30.2	30.5	27.9	28.1	28.1	27.9	28.1	26.6	26.9	27.1	27.2
F10	13.5	13.7	13.6	13.7	13.7	13.6	13.7	13.0	13.2	13.3	13.4
F11	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
F12	14.2	14.4	10.2	10.2	10.3	10.2	10.2	7.3	7.4	7.4	7.5
F13	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
F14	1.0	1.0	0.7	0.7	0.7	0.7	0.7	0.5	0.5	0.5	0.5
F15	0.5	0.5	0.7	0.7	0.7	0.7	0.7	0.4	0.4	0.5	0.5
F16	2.4	2.4	2.5	2.6	2.6	2.5	2.6	0.1	0.1	0.1	0.1
F17	5.1	5.1	4.9	5.0	5.0	4.9	5.0	4.9	5.0	5.0	5.0
F18	4.1	4.2	5.4	5.5	5.5	5.4	5.5	8.5	8.5	8.6	8.7
F19	1.3	1.3	0.9	0.9	0.9	0.9	0.9	0.7	0.7	0.7	0.7
F20	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
F21	2.2	2.2	1.7	1.7	1.7	1.7	1.7	1.6	1.6	1.6	1.6
F22	0.6	0.6	0.5	0.5	0.5	0.5	0.5	0.4	0.4	0.4	0.4
F23	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F24	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2
F25	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.5	0.5	0.5	0.5
F26	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2
F27	10.4	10.5	11.5	11.6	11.6	11.5	11.6	12.2	12.4	12.5	12.6
F30	9.9	10.0	7.7	7.7	7.7	7.7	7.7	9.5	9.7	9.7	9.8
F31	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.4	1.4	1.5	1.5
F32	4.8	4.8	4.0	4.1	4.1	4.0	4.1	3.3	3.4	3.4	3.4
G	6.6	6.7	7.0	7.1	7.1	7.0	7.1	7.1	7.2	7.2	7.3
H	4.7	4.8	7.0	7.1	7.1	7.0	7.1	11.7	11.9	11.9	12.0
I	10.8	10.9	10.9	11.0	11.0	10.9	11.0	11.2	11.3	11.4	11.5
J	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.7	0.7	0.7	0.7
K	7.2	7.2	9.4	9.5	9.5	9.4	9.5	10.9	11.0	11.1	11.2
L	49.1	49.6	50.9	51.3	51.4	51.0	51.4	52.9	53.5	53.9	54.2
M	18.8	19.0	19.8	20.0	20.0	19.9	20.0	17.9	18.1	18.3	18.4
N	8.2	8.3	9.6	9.7	9.7	9.7	9.7	11.1	11.3	11.4	11.4
O	6.2	6.2	6.2	6.3	6.3	6.2	6.3	6.7	6.8	6.8	6.9
Q	22.5	22.7	25.7	25.9	25.9	25.7	25.9	24.1	24.3	24.5	24.7

Table 6: intermediate expenditure for each industry (1 trillion yen)

Industry	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
A	12.3	12.3	10.9	10.9	10.9	10.9	10.9	10.7	10.7	10.7	10.7
B	1.7	1.7	1.1	1.1	1.1	1.1	1.1	0.8	0.8	0.8	0.8
C	2.6	2.6	2.4	2.4	2.4	2.4	2.4	2.1	2.1	2.1	2.1
D	9.3	9.3	11.7	11.7	11.7	11.7	11.7	18.2	18.2	18.2	18.2
E	7.6	7.6	8.4	8.4	8.4	8.4	8.4	8.5	8.5	8.5	8.5
F09	12.5	12.5	13.2	13.2	13.2	13.2	13.2	12.8	12.8	12.8	12.8
F10	4.8	4.8	4.7	4.7	4.7	4.7	4.7	4.8	4.8	4.8	4.8
F11	4.3	4.3	2.8	2.8	2.8	2.8	2.8	2.0	2.0	2.0	2.0
F12	2.7	2.7	1.9	1.9	1.9	1.9	1.9	2.0	2.0	2.0	2.0
F13	7.0	7.0	5.4	5.4	5.4	5.4	5.4	4.4	4.4	4.4	4.4
F14	3.2	3.2	2.9	2.9	2.9	2.9	2.9	3.5	3.5	3.5	3.5
F15	10.3	10.3	9.5	9.5	9.5	9.5	9.5	9.3	9.3	9.3	9.3
F16	11.4	11.4	11.3	11.3	11.3	11.3	11.3	6.1	6.1	6.1	6.1
F17	26.1	26.1	26.9	26.9	26.9	26.9	26.9	28.8	28.8	28.8	28.8
F18	11.3	11.3	13.0	13.0	13.0	13.0	13.0	15.7	15.7	15.7	15.7
F19	11.1	11.1	11.0	11.0	11.0	11.0	11.0	11.8	11.8	11.8	11.8
F20	3.1	3.1	3.0	3.0	3.0	3.0	3.0	2.8	2.8	2.8	2.8
F21	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
F22	12.5	12.5	10.4	10.4	10.4	10.4	10.4	8.7	8.7	8.7	8.7
F23	21.6	21.6	18.2	18.2	18.2	18.2	18.2	26.2	26.2	26.2	26.2
F24	8.6	8.6	8.0	8.0	8.0	8.0	8.0	9.8	9.8	9.8	9.8
F25	17.1	17.1	14.8	14.8	14.8	14.8	14.8	14.2	14.2	14.2	14.2
F26	10.0	10.0	10.3	10.3	10.3	10.3	10.3	11.0	11.0	11.0	11.0
F27	23.2	23.2	25.4	25.4	25.4	25.4	25.4	22.8	22.8	22.8	22.8
F30	21.9	21.9	22.1	22.1	22.1	22.1	22.1	28.7	28.7	28.7	28.7
F31	1.5	1.5	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6
F32	3.4	3.4	4.6	4.6	4.6	4.6	4.6	3.6	3.6	3.6	3.6
G	15.0	15.0	15.5	15.5	15.5	15.5	15.5	15.1	15.1	15.1	15.1
H	9.0	9.0	13.4	13.4	13.4	13.4	13.4	26.1	26.1	26.1	26.1
I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
J	0.1	0.1	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
K	27.9	27.9	26.5	26.5	26.5	26.5	26.5	26.9	26.9	26.9	26.9
L	10.4	10.4	8.9	8.9	8.9	8.9	8.9	8.1	8.1	8.1	8.1
M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N	1.0	1.0	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9
O	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4
Q	62.1	62.1	74.1	74.1	74.1	74.1	74.1	69.7	69.7	69.7	69.7

References

Alvarez, I. and J. Molero (2005) “Technology and the generation of international knowledge spillovers: An application to Spanish manufacturing firms,” *Research Policy*, Vol. 34, pp. 1440–1452.

Arellano, M. and S. Bond (1991) “Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations,” *Review of Economic Studies*, Vol. 58, pp. 277–297.

Baldwin, R., R. Forslid, P. Martin, G. Ottaviano, and F. Robert-Nicoud (2003) *Economic Geography and Public Policy*: Princeton University Press.

Blundell, R. and S. Bond (1998) “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics*, Vol. 87, pp. 115–143.

Bode, E. (2004) “Productivity Effects of Agglomeration Externalities.”

Brakman, S., H. Garretsen, J. Gorter, A. V. D. Horst, and M. Schramm (2009) “New Economic Geography, Empirics, and Regional Policy.”

Brülhart, M. and N. A. Mathys (2008) “Sectoral agglomeration economies in a panel of European regions,” *Regional Science and Urban Economics*, Vol. 38, No. 4, pp. 348–362.

Ciccone, A. (2002) “Agglomeration effects in Europe,” *European Economic Review*, Vol. 46, pp. 213–227.

Ciccone, A. and R. E. Hall (1996) “Productivity and the Density of Economic Activity,” *American Economic Review*, Vol. 86, pp. 54–70.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2008) “Estimating Agglomeration Economies with History, Geology, and Worker Effects,” *CEPR Discussion Paper*, No. DP6728.

Crozet, M. (2004) “Do migrants follow market potentials? An estimation of a new economic geography model,” *Journal of Economic Geography*, Vol. 4, pp. 439–458.

Davis, D. R. and D. E. Weinstein (2008) “A Search for Multiple Equilibria in Urban Industrial Structure,” *Journal of Regional Science*, Vol. 48, No. 1, pp. 29–65.

Dixit, A. K. and J. E. Stiglitz (1977) “Monopolistic Competition and Optimum Product Diversity,” *American Economic Review*, Vol. 67, No. 3, pp. 297–308.

Fosfuri, A. and T. Ronde (2004) “High-tech clusters, technology spillovers, and trade secret laws,” *International Journal of Industrial Organization*, Vol. 22, pp. 45–65.

Fujita, M., P. Krugman, and A. Venables (1999) *The Spatial Economy: Cities, Refions and International Trade*: MIT Press.

Fujita, M. and J. F. Thisse (2002) *Economics of Agglomeration*: Cambridge University.

Hanson, G. H. (2005) “Market Potential, Increasing Returns, and Geographic Concentration,” *Journal of International Economics*, Vol. 67, No. 1, pp. 1–24.

Henderson, J. V. (2007) "Understanding knowledge spillovers," *Regional Science and Urban Economics*, Vol. 37, pp. 497–508.

Klette, T. J. and Z. Griliches (1996) "The inconsistency of common scale estimators when output prices are unobserved and endogenous," *Journal of Applied Econometrics*, Vol. 11, No. 4, pp. 343–361, July.

Krugman, P. R. (1980) "Scale economies, product differentiation, and the pattern of trage," *American Economic Review*, Vol. 70, pp. 950–959.

Levinsohn, J. and M. Melitz (2002) "Productivity in a differentiated products market equilibrium," *mimeo*, No. 1996, pp. 1–16.

Mion, G. (2004) "Spatial externalities and empirical analysis: the case of Italy," *Journal of Urban Economics*, Vol. 56, No. 1, pp. 97–118, July.

Monjon, S. and P. Waelbroeck (2003) "Assessing spillovers from universities to firms: evidence from French firm-level data," *International Journal of Industrial Organization*, Vol. 21, pp. 1255–1270.

Ornaghi, C. (2006) "Spillovers in product and process innovation: Evidence from manufacturing firms," *International Journal of Industrial Organization*, Vol. 24, pp. 349–380.

Pons, J., E. Paluzie, J. Silvestre, and D. A. Tirado (2007) "Testing The New Economic Geography: Migrations And Industrial Agglomerations In Spain," *Journal of Regional Science*, Vol. 47, No. 2, pp. 289–313.

Redding, S. J. (2010) "The Empirics Of New Economic Geography," *Journal of Regional Science*, Vol. 50, pp. 297–311.

Tveteras, R. and G. E. Battese (2006) "Agglomeration Externalities, Productivity, and Technical Inefficiency," *Journal of Regional Science*, Vol. 46, No. 4, pp. 605–625, October.

Wooldridge, J. M. (2002) "Econometric Analysis of Cross Section and Panel Data," Vol. 58, No. 2, p. 752.