The Demand- and Supply-Side Spatial Spillovers in the Food Processing Industry in Korea:

An Empirical Evidence from Both Local Level and Individual Firm Level

by

Donghwan An, Kwansoo Kim, and OhSang Kwon

Department of Agricultural Economics and Rural Development, Seoul National University, Korea.

Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Denver, Colorado, August 1-4, 2004

ABSTRACT

From the perspective of the food system, the overall food manufacturing or processing industry is fundamentally connected both with agricultural production in rural areas and consumption demand usually concentrated in urban areas. Using the local government level and individual firm level data for the food manufacturing industry in Korea, this paper investigated the agglomeration and spillover effects in this industry. This study found that there exist significant productivity differentials over space in food processing industry. This paper also found the evidences of agglomeration economies; the place where the size of population is large performs better. The results showed some evidences of spillover effects; negative externalities from congestions of neighbors and positive spillovers of the increasing accessibility to material input producers in the neighboring regions.

Key words : Food Manufacturing Industry, Agglomeration Economies, Spillover Effects

I. Introduction

As seen from the economics literature on agglomeration and economic geography, various factors including scale economies in the production of intermediate inputs and knowledge spillovers motivate a firm to be located in a region where other similar types of firms are in close proximity. However, the literature on the agglomeration effects on location decision and economic performance remains sparse from the perspective of individual industrial sector - especially in an empirical context.[1]

In this study, we intend to investigate the spillover effects associated with the demand-driven and supply-side externalities in the food manufacturing industry in Korea. The food manufacturing (or processing) industry demands farm products and supplies food. That is, this sector is fundamentally connected both with agricultural production in rural areas, and consumption demand usually concentrated in urban areas (Cohen and Paul, 2001). Particularly, in Korea, the investigation of demand-driven spillover effects on the firms' location decision in this industry is of interest for two reasons. First, as one of the major importers of agricultural products, Korea heavily depends on imports to meet the raw material input demands in this industry. And hence, the demand-side spillover effects across food manufacturing firms are particularly of interest in the context of Korea being an importing country. Second, in Korea, promoting the food processing sector that uses local farm products as inputs has been an important rural development policy measure in the sense of increasing farm income and improving rural economy.

This study attempts to provide useful information on these issues by developing an empirical framework of measuring and identifying these spillover effects. Both aggregated data (at the local government level) and individual firm data are used for this study. We first develop a model allowing us to recover the information on the performance of this industry. We incorporate the measures of demand- and supply-side agglomeration factors into our framework and test whether these spillover effects are significant or not. In doing this, we use a frontier production function approach. In the frontier literature, the difference in productivity of performance of firms is termed "inefficiency." Many of the previous studies on regional productivity issues generally assume that there exists no technical inefficiency. However, taking this short-fall into consideration, many researchers have shown the existence of significant differences in technical efficiency among regions (Beeson and Husted, 1989; Chambers et al., 1996a; Domazlicky and Weber, 1997; Kim, 1997). In this line of research, the frontier production function approach has been increasingly used for regional productivity analysis in years (Puig-Junoy, 2001; Färe et al., 2001). To represent the production technology, we use a stochastic frontier model developed by Battese and Coelli (1995). In this paper, we define agglomeration economies as the degree of externalities within a ‘given area’, measured by the conventional agglomeration factors (i.e., population size and own-industry employment). We also define spatial spillover effects as the degree of externalities captured by agglomeration factors of ‘neighboring areas’.

A number of studies conducting empirical analysis on the investigation of the performance of Korean manufacturing industry have been done (e.g., Kim, 1997; Koo and Kim, 1999; Lee, 2000; Park and Cho, 2001; Henderson et al., 2001; An et al., 2003). However, most of them rely on the conventional approach ignoring technical inefficiency. Relatively little attention has been paid to the investigation of the productivity issues in the food manufacturing industry. In addition, as far as we know, no empirical studies have used individual firm data in Korea.

The structure of this paper is as follows. The next section discusses the data and the estimation model employed for this study. The subsequent section will provide the estimation results. Finally, findings are summarized with some concluding remarks.

II. A Model

The effects of spatial proximity and concentration on firms' performance and their location decision have been recognized, at least, since early 1900s (e.g., Marshall, 1920; Hoover, 1937). The idea that firms can reduce production costs by locating closely with other firms stimulated a huge literature on agglomeration economies. Agglomeration economies, a kind of positive externalities in production, occur when economic activities are located within one area, which is one of the central concepts among regional scientists.

Agglomeration economies are usually categorized into two types; localization economies and urbanization economies. Localization economies occur if firms in a particular industry located in one place can reduce production cost as the size of the industry increases in the region. Localization economies result from the factors that are 'external to the individual firm but internal to the local industry'. O'Sullivan (2000) identified three major factors representing localization economies; scale economies in the provision of intermediate inputs, labor market pooling, and knowledge spillovers. Urbanization economies occur if the production cost of an individual firm decreases as the scale of the entire economy of a particular region increases. Therefore, urbanization economies are 'external to local industry but internal to the local economy'.

Many of the early studies estimated production functions using cross-section data for urban manufacturing industries. It was generally assumed that productivity differentials among regions were captured by Hicks-neutral shifts of an aggregate regional production function. The typical version of the production function approach estimated the following production function, f(x), with Hicks-neutral external factor g(A) ignoring the technical inefficiency.

(1) $y = g(A)f(x)$

Here, y, x, and A are output, inputs, and the factors related to the external economies or agglomeration economies, respectively. In this approach, the agglomeration factors A are considered as shift factors and included in the production function in the same manner as the inputs. Two important issues in the estimation of equation (1) are the specification of g(A) and the assumption of Hicks-neutral productivity shift. Most frequently used measures to representing localization economies and urbanization economies are own-industry employment and population, respectively.[2]

This study use the frontier function approach rather than including agglomeration factors into the production function as in equation (1). The empirical literature that focuses on frontier production has used two broadly defined approaches; the nonparametric programming approach known as data envelopment analysis (Charnes et al. 1978); the parametric stochastic approach (Aigner et al. 1977). Although these two approaches are based on similar theoretical foundation, they often produce different empirical results. This paper employs the stochastic frontier model suggested by Battese and Coelli (1993, 1995) and Huang and Liu (1994).

The stochastic frontier approach allows the deviations from the frontier technology resulting from both technical inefficiency and random factors. Stochastic frontier production function is defined in equation (2) in which an error term e is decomposed into two terms u and v.

$$(2) \qquad y = f(x) + e,\ e = u - v,\ v \geq 0$$

Here, u is a symmetrically distributed noise (random error) term and v is a non-negative inefficiency term. Under this formulation, technical inefficiency can be estimated as conditional expectation $E[v \mid e]$..

In particular, a stochastic frontier model developed by Battese and Coelli (1993, 1995) and Huang and Liu (1994) allows heteroschedasticity in v, which should preferably be flexible. Their models allow the inefficiency effects to be a function of a set of explanatory variables the parameters of which are estimated simultaneously with the frontier production function f(x). This is one of the most advantageous features of their models because it allows us to investigate the determinants of production efficiency in one step.

Following their specification of frontier production function, equation (2) is redefined as follows:

(3) $$y = f(x) + u - v, \quad v = dZ + w.$$

Here, u is assumed to be independently and identically distributed random errors that have normal distribution with mean zero and variance σ_u^2, i.e., $u \sim N(0, \sigma_u^2)$; v is non-negative inefficiency term truncated at zero and it is assumed to be independently and identically distributed, i.e., $v \sim N^+(dZ, \sigma_v^2)$; Z is a vector of explanatory variables associated with production inefficiency; w is a random variable defined by the truncation of normal distribution with mean zero and variance σ_v^2, i.e., $w \sim N^+ (0, \sigma_v^2)$.

Technical efficiency effect model, defined as a linear function of Z, can be estimated simultaneously with frontier production function f(x) using maximum likelihood estimation method (Huang and Liu, 1994; Battese and Coelli, 1993, 1995). That is, the coefficients of production function, f(x), and technical efficiency effect model, dZ, are simultaneously estimated. Technical efficiency can be predicted by $E(-v)=E(-dZ-w)$.

III. Data and Estimation Results

1. Data

The data used for this study are obtained from two sources. First, the aggregated data at local government level are obtained from the Mining and Manufacturing Survey for the period of 1991-2001, conducted by Korea National Statistics Office. This survey provides input-output relationships of individual firms in manufacturing industry. Unfortunately, however, the geographical information of individual firms is not disclosed for public use. This forces us to use only aggregated data at local government level. This data set includes 230 local governments among the total of 234 local governments, and the total number of observations is 2,410. Second, the accounting data for 248 individual firms that are listed in the stock market are obtained from Korea Investors Service, a leading credit rating agency in Korea. This data set on individual firm's accounting information is available for the period of 1997-2001, and the total number of observations is 924. These two data sets together are used for the estimation of production function f(x) in equation (3).

In specifying the production function f(x), we use total value of product (y) as an output measure. As inputs, four factors are included; labor (L), capital (C), raw materials (R), and miscellaneous inputs such as fuel and electricity (O). For labor input measure, we use total employment. Real capital stock, measured as the beginning year capital stock (total value of tangible fixed assets), is included as capital input. Explanatory variables associated with production inefficiency (Z) include time, population, own-industry employment size and farmland of a region and those of neighboring regions. Time variable (T) is included to allow efficiency to vary over time. While the degree of urbanization economies can be measured by the size of population at a 'given area' (POP), total population of 'neighboring areas' (WPOP) captures the demand-side spatial spillover effects. Similarly, the effects of localization economies can be measured by own-industry employment size within a region (WORK), while that of neighboring region (WWORK) captures the localization spillovers over space. In this paper, we include farmland area as another localization factor capturing the supply capacity of material inputs for this industry. We define farmland within a region (FLAND) as the supply-side agglomeration factor in food processing sector and farmland of neighboring region (WFLAND) as the measure of supply-side spillovers. We also include a distance measure (DIST) measuring the distance from the nearest metro city among five major metropolitan areas (Seoul, Busan, Daejun, Daegu, Gwangju) to capture the proximity or accessibility to a place with high demand density and urban center. It is expected that distance to the major metropolitan areas may capture the impact of demand density that seems particularly important (operative) in this industry. The factors capturing spatial spillover effects are constructed using a spatial contiguity matrix. That is, they are measured by the weighted sum of neighboring areas' agglomeration factor. A spatial contiguity matrix W is constructed as follows. If region i and j has common boundary, the element of this matrix, w_{ij}, is 1 (otherwise, w_{ij} is zero). After making transformation that converts the matrix W to have row-sums of unity and multiplying the agglomeration factors at each region produce the weighted sum of each agglomeration factors capturing 'spatial spillovers'. For the analysis using aggregated data, we add average firm size of each region (SIZE) to control for scale effects. All efficiency factors except distance and average size variables are expressed in logarithm.

Table 1. Summary Statistics of Output, Inputs and Inefficiency Factors

Variables	Individual Firm Data: Model I (N=924)		Aggregated Data: Model II (N=2410)	
	Mean	Standard Deviation	Mean	Standard Deviation
Output (Million Won)*	141,392	294,355	129,316	199,289
Labor (Person)	539	1,010	511	600
Capital (Million Won)*	63,573	168,890	53,784	89,941
Raw material (Million Won)*	56,232	127,545	72,867	117,331
Other inputs (Million Won)*	14,470	36,823	5,484	9,178
Population (1000 Person)	270	177	205	177
Industry employment (Person)	1,853	1,189	863	981
Farmland (ha)	10,342	8,754	8,874	8,067
Neighborhood population (Person)	226	116	200	121
Neighborhood industry employment (Person)	1,201	652	847	574
Neighborhood farmland (ha)	10,449	6,974	9,357	7,070
Distance from metro city (km)	47	30	51	46
Average firm size(Person/Firm)	-	-	32	27

* These are in real terms deflated by GDP deflator (1995=100).

2. Estimation Results

We estimate equation (3) using two different data sources, individual firm level (Model I) and local level data (Model II). Table 2 shows the estimation results of two models. The estimates of σ^2 and $\sigma_v^2/(\sigma_u^2 + \sigma_v^2)$ from both models are estimated to be statistically significant at 1%. This shows our model considering technical inefficiency is appropriate for the data set.

The estimation results from two models seem to be consistent except the coefficients associated with supply-side spillover effects measured by the farmland area of the region (lnFland and lnWFLAND). In the inefficiency model, the negative sign indicates that technical inefficiency (or technical efficiency) decreases (or increases) as the associated explanatory variable increases.

The coefficient of time variable (T) is estimated to be negative and significant in Model II, but not significant in Model I. This shows that the food manufacturing industry in Korea improves its performance during 1990s at the aggregated level. At the firm level, the coefficient of time variable is positive but insignificant.

The coefficient of own-region population (lnPOP) is estimated as positive in both models. This provides empirical evidences of urbanization economies in this industry meaning that the food processing industry in more populated region performs better than less populated area, although these effects are estimated not to be significant in the firm level model, Model I.

On the contrary, the coefficient associated with population size of neighboring regions (lnWPOP) is positive in both models, although it is estimated to be insignificant in Model I. This means that the population size of neighboring regions is negatively related with the performance of the firms in this industry, i.e. negative demand-side spatial spillover effects. This might also imply that the benefits of urbanization economies from neighboring highly populated regions may not be large enough to compensate the negative spillover effects of the congestion.

The coefficient associated with own-industry employment in a region is estimated to be positive and significant in both models. This is unexpected, because own-industry employment is commonly used to measure the localization economies. Our estimation results indicate that the firms locating in a region where the food processing industry is concentrated performs worse than other regions. This might imply that the level of competition in this industry is very high for both input and output markets and this leads to relatively higher level of production costs.

Unlike localization economies within a region, the coefficient of own-industry employment of neighboring regions is estimated as negative and statistically significant for both models. This means that spatial spillovers of localization economies from neighboring regions are positively related with the performance of the food processing firms.

Table 2 shows that the estimation results associated with the supply-side agglomeration factor are very contradictory in two models. The farmland area in a region (lnFLAND) is included as a proxy measure of accessibility to the material input producers (agricultural products) indicating the supply-side agglomeration factor. The positive (negative) sign of this variable means that the firms located in the region where supply capacity of material input (i.e. agricultural products) is higher are likely to perform poor (better). The coefficient associated with farmland area in a region is estimated to be negative and significant in Model I, while it is positive and significant in Model II. The estimation result of Model I indicates that the input-supply capacity is positively related with the performance of the individual firms, while Model II says it is negatively related with the performance of this industry at the local level.

The estimation results associated with supply-side spillover effects is also mixed in two models. The coefficient associated with the farmland area in the neighboring region (lnWFLAND) is estimated to be positive and significant in Model I, while negative and insignificant in Model II. This indicates that the supply-side spillover effects are negative or at most negligible in Korea.

Several explanations seem plausible for this mixed effect of supply-side agglomeration factor. One is that the food manufacturing industry in Korea does not depend much on domestic supply of agricultural products for its raw materials. It seems that the region where the farmland area is relatively large and agriculture is a major industry is not likely to provide a good environment for manufacturing firms in Korea. And also, the limitations on the data set used for this study are also attributable to the results. We observed the individual firms used for Model I only in 106 regions (local governments), which is much smaller than that of Model II (230). As seen in Table 1, the distribution of Z variables is different between two data sets for Model I and Model II.

Considering the very high concentration of economic activities into metropolitan areas in Korea, we include a distance measure (DIST) defined by the distance from the nearest metropolitan areas in order to measure the agglomeration spillovers from large metropolitan areas with high demand density. In both models, the coefficient associated with the distance measure is estimated to be positive and significant. This indicates that, as the distance from the major metropolitan areas increases, technical inefficiency increases.

Finally, we include average firm size (size) for Model II to control the scale effects, because Model II uses aggregated data at the local level. Our estimation results indicate that the scale of individual firms is positively related with the performance of food manufacturing industry at the local level.

Table 2. Estimation Results of Stochastic Production Frontier

Variables		Model I: Firm level model	Model II: Local government level model
Production function, f(x)	Constant	1.73E+01 (1.92E+00)*	2.33E+00 (1.30E-01)*
	lnL	-1.18E-01 (1.69E-01)	3.08E-01 (4.58E-02*
	lnC	-1.29E+00 (2.60E-01)*	-1.74E-01 (2.75E-02)*
	lnR	1.06E-01 (1.12E-02)*	4.50E-01 (2.81E-02)*
	lnO	-7.03E-02 (1.49E-02)*	4.34E-01 (3.52E-02)*
	lnLlnL	1.07E-01 (5.93E-02)***	-1.04E-01 (1.61E-02)*
	lnClnC	8.37E-02 (2.14E-02)*	7.33E-02 (6.89E-03)*
	lnRlnR	6.65E-02 (7.10E-03)*	1.65E-01 (7.46E-03)*
	lnOlnO	2.19E-02 (8.88E-03)*	-3.67E-02 (1.37E-02)*
	lnLlnC	-1.25E-03 (1.97E-04)*	2.19E-03 (1.01E-02
	lnLlnR	4.24E-04 (2.86E-04)	-1.00E-02 (8.49E-03)
	lnLlnO	4.47E-04 (2.86E-04)	6.52E-02 (6.52E-03)*
	lnClnR	-2.08E-05 (8.02E-05)	-7.73E-02 (6.45E-03)*
	lnClnO	2.64E-04 (7.00E-05)*	3.84E-02 (8.70E-03)*
	lnRlnO	-3.73E-04 (4.48E-05)*	-7.73E-02 (9.85E-03)*
	T	3.31E-01 (3.83E-01)	-7.67E-02 (9.66E-03)*
	TT	1.14E-02 (1.82E-02)	7.68E-03 (6.66E-04)*
	TlnL	2.61E-02 (1.45E-02)***	-1.25E-02 (3.39E-03)*
	TlnC	-1.97E-03 (7.32E-04)*	7.99E-03 (2.54E-03)*
	TlnR	-3.10E-03 (3.93E-03)	-9.22E-03 (1.92E-03)*
	TlnO	-3.93E-04 (4.58E-03)	1.07E-02 (2.58E-03)*
Inefficiency factors, Z	Constant	-1.17E-01 (6.68E-01)	1.93E-02 (1.86E-02)
	T	1.47E-01 (3.71E-01)	-1.93E-03 (1.44E-04)*
	lnPOP	-3.28E-02 (2.46E-02)	-9.85E-03 (9.10E-04)*
	lnWORK	7.60E-02 (2.53E-02)*	1.22E-02 (2.66E-03)*
	lnFLAND	-2.89E-02 (1.16E-02)*	4.95E-03 (1.41E-03)*
	lnWPOP	4.06E-02 (4.81E-02)	3.35E-03 (1.44E-03)**
	lnWWORK	-1.34E-01 (3.50E-02)*	-5.04E-03 (2.09E-03)**
	lnWFLAND	9.93E-02 (2.77E-02)*	-2.46E-03 (1.52E-03)
	DIST	9.37E-04 (6.87E-04)	3.48E-04 (6.40E-05)*
	SIZE	- -	-9.65E-04 (9.64E-05)*
σ^2		2.01E-01 (9.59E-03)*	2.79E-02 (8.10E-04)*
$\sigma_v^2/(\sigma_v^2+\sigma_v^2)$		4.45E-02 (2.17E-02)**	2.15E-03 (9.05E-05)*
Log-likelihood		-5.66E+02 (3.66E+01)**1)	8.76E+02 (6.93E+01)** 1)

* significant at 1%, ** significant at 5%, *** significant at 10%

1) Test statistics for ‘No technical inefficiency’ hypothesis.

V. Summary and Conclusion

This study investigates the spillover effects associated with the demand- and supply-side agglomeration economies in the food manufacturing industry in Korea. Both aggregated data and individual firm data are used and the estimation results of stochastic frontier production function are compared.

Our estimation results show that (i) there exist significant productivity differentials over space in food processing industry; (ii) the firms located in the place where the size of population is large performs better, while there also exist the negative externalities from the congested neighbors at the local industry level; (iii) the local own-industry employment within region is negatively relative with the performance of firms in this industry, while there exist positive spatial spillovers from the own industry agglomeration in the neighboring region; (iv) the effects of increasing accessibility to material input producers on the performance of food manufacturing firms are controversial depending on the data set used.

This study shows that the analysis unit matters for the agglomeration studies and the spatial spillover researches. The implications from this analysis are limited by using different data sources. In particular, our model using individual firms do not consider the small firms which are not listed in the stock market but consists of non-negligible parts of food manufacturing industry. The variables included for representing agglomeration and spillover factors are also too simplified. For the individual firm level model, this study does not take into account the individual firm’s characteristics affecting its performance, and hence omission errors might be critical.

References

Battese, G. E. and T. J. Coelli. “A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data”, *Empirical Economics* 20 (1995): 325-332.

Battese, G. E. and T. J. Coelli. “A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects”, Working Papers in Econometrics and Applied Statistics No. 69 (1993), University of New England.

Beeson, P. and S. Husted. “Patterns and Determinants of Productive Efficiency in State Manufacturing”, *Journal of Regional Science* 29 (1989): 15-28.

Cohen, J. P. and C. J. M. Paul. “Agglomeration Economies and Industry Location Decisions: The Impact of Vertical and Horizontal Spillovers”, Working Paper No. 01-004 (2001), Department of Agricultural and Resource Economics, University of California Davis.

Coelli, T. “A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation”, Centre for Efficiency and Productivity Analysis Working Paper96/07 (1996), University of New England, Australia.

Domazlicky, B. R. and W. L. Weber. "Determinants of Total Factor Productivity, Technological Change, and Efficiency Differentials Among States, 1977-86," *Review of Regional Studies* 28 (1998): 19-33.

Domazlicky, B. R. and W. L. Weber. "Total Factor Productivity in the Contiguous United States, 1977-86", *Journal of Regional Science* 37 (1997): 213-233.

Eberts, R. W. and D. P. McMillen. "Agglomeration Economies and Urban Public Infrastructure", in *Handbook of Regional and Urban Economics*, Vol. III, ed. by P. Cheshire and E. S. Mills, Elsevier Science B. V.:Amsterdam, 1999.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang. "Productivity Growth, Technical Progress, and Efficiency Change in Industrial Countries", *The American Economic Review* 84 (1994): 66-83.

Henderson, J. V. "Marshall's Scale Economies", NBER Working Paper 7358 (1999), National Bureau of Economic Research.

Henderson, J. V. "Efficiency of Resource Usage and City Size", *Journal of Urban Economics* 19 (1986): 47-70.

Henderson, V., T. Lee, and Y. J. Lee. "Scale Externalities in Korea", *Journal of Urban Economics* 49 (2001): 479-504.

Huang, C. J. and J. Liu. “Estimation of a Non-Neutral Stochastic Frontier Production Function”, *The Journal of Productivity Analysis* 5 (1994): 171-180.

Kim, S. J. *Productivity of Cities*, Ashgate Publishing Company, 1997.

Lee, B. S. “Determinants of Productivity of Manufacturing Industries in Seoul Metropolitan Area and Its Suburbs”, *Journal of The Korean Economic Association* 48 (2000): 291-322 (in Korean).

O'Sullivan A. *Urban Economics*, 4th edition, McGraw-Hill Companies, 2000.

Park, H. S. and K. Y. Cho. *Estimation of Externalities from Concentration to Seoul Metropolitan Area and Its Suburb*, Research Report, Kyonggi Research Institute, 2001 (in Korean).

Puig-Junoy, J. “Technical Inefficiency and Public Capital in U.S. States: A Stochastic Frontier Approach”, *Journal of Regional Science* 41 (2001): 75-96.

Richardson, H. W. “Economies and Diseconomies of Agglomeration, in Urban Agglomeration and Economic Growth”, ed. by H. Giersch, Berlin: Springer-Verlag, 1995.

---

[1] See Henderson (1999), Eberts and McMillen (1999), and Richardson (1995), etc.

[2] One of most frequently used specification for g(A) is $g(A)=\exp(r/L)N^e$ suggested by Henderson (1986), where L and N are own-industry employment and population in the region. Here, the output elasticities of localization and urbanization are measured by $-r/L$ and e, respectively.