A Spatial Econometric Model of the Korean Economy

Doleswar Bhandari, Ph.D.
University of New Mexico
Bureau of Business and Economic Research
303 Girard Blvd, STE 116
1 University of New Mexico
Albuquerque, New Mexico 87131
Tel. 505-277-7067
Fax 505-277-7066
bhandar1@unm.edu

Tom Johnson, Ph.D.
215 Middlebush Hall
Community Policy Analysis Center
University of Missouri-Columbia
Columbia, MO 65201
573-882-2157
JohnsonTG@missouri.edu

Dennis P. Robinson, Ph.D.
215 Middlebush Hall
Community Policy Analysis Center
University of Missouri-Columbia
Columbia, MO 65201
573-882-7818
Robinsonden@missouri.edu

A SPATIAL ECONOMETRIC MODEL OF THE KOREAN ECONOMY

ABSTRACT

Using Korean regional data a spatial simultaneous equation model was developed and estimated using a generalized spatial three-stage least squares procedure. This model contains 10 equations for local finance, the labor market, and the housing market -local revenue, local expenditures, housing units, population, economically active population, number of students, in- and out-commuting, number of firms, and employment in non-basic sectors. Employment and economic development expenditures are the main drivers of the model. A significant cross-county and cross-equation spillover effect is estimated. Reduced form estimates were derived from structural equations, and it is found that additional employment opportunities generated in a certain county will have a positive impact on all sectors of the economy, including local finance, the housing market, demography, and the labor market. The spatial spillover effect is estimated on neighboring counties resulted from the employment opportunities created in a residence county. It appears that the estimated parameters tend to be sensitive to the specification of weight matrices. The model was validated based on forecasting accuracy of in-sample data using mean absolute percentage error.

1. Introduction

The development of econometric models for economic impact analysis and economic forecasting has been of continuous interest to regional economists. In this paper such a model is applied to the Korean regional economy. Regions within a country are open economies experiencing extensive inter-area spillovers. For example, employment change in one region may affect the population, commuting pattern, and demand for public services in nearby area. To account for the interregional spillovers, a 'complete' version of spatial model was developed for the Korean economy using generalized spatial three-stage least squares (GS3SLS) procedure.

This model is somewhat similar to other fiscal impact models developed in the United States based on county-level units of analysis, the most common of which are the Show Me Model (Johnson and Scott, 2006), the Virginia Impact Projection Model (Johnson, 1991), the Iowa Economic/Fiscal Impact Modeling System (Swenson and Otto, 1998), the Idaho model (Cooke and Fox, 1994), an Integrated Economic Impact and Simulation Model for Wisconsin Counties (Shields, 1998), and the Small Area Fiscal Estimation Simulator for Texas (Evans and Stallmann, 2006). One of the key elements in these models is employment as an engine of economic growth and change at the local level.

Using Korean data for 2005, a model is estimated using a GS3SLS procedure developed by Kelejian and Prucha (2004). The model contains equations of local revenues, local expenditures, total housing units, population, total students, in- and out-communing, number of firms, and non-basic employment. A simultaneous system of spatially interrelated cross-section equations containing spatial lags in both the dependent variables and their error terms is considered. Feedback effects due to both the simultaneous relationships between equations and the spatial linkages are also investigated. To validate the model, a non-spatial model (NS3SLS), simultaneous spatial lag model (SL3SLS), spatial error model (SE3SLS), and spatial lag and spatial error model (SLE3SLS) are estimated. Using a measure of forecasting accuracy (mean absolute percentage error [MAPE]), the performance of the models is investigated.

The rest of the paper is organized as follows. Section 2 outlines the objectives of this research; Section 3 discusses the data and data sources; and Section 4 outlines the spatial model estimation procedure. Sections 5 and 6 discuss the model specification and results, respectively, and Section 7 summarizes and concludes.

2. Research Objectives

The primary objective of this study is to build a rigorous econometric model for the South Korean economy that can be applied to alternative policy scenarios. The specific objectives are:

1. to develop and estimate a GS3SLS model for local regions in Korea,
2. to estimate the interregional spillover effect of the labor market and the local public finance and housing market, and
3. to perform policy analysis based on the reduced form estimates.

3. Data

This empirical work focuses on the relationships and spatial interactions among labor market, the local public finance market, and the housing market. Although Korean regions are comparatively centralized entities, it is assumed that these entities are independent decision-making bodies. Administrative divisions of South Korea are divided into 1 Special City (teukbyeolsi), 6 Metropolitan Cities (gwangyeoksi), and 9 Provinces (do). Based on population, these are further subdivided into a variety of smaller entities, including cities (si), counties (gun), wards (gu), towns (eup), districts (myeon), neighbourhoods (dong) and villages (ri). The data used in this study come from 172 local government units that consist of 7 metropolitan cities, 77 cities, and 88 counties. Among the 172 regions, 140 are rural, with a population of less than 270,000, and 32 are urban, with a population of more than 270,000. Thus, the analysis includes all 172 regions in Korea with population over 8,000.

All the data used to conduct this study are secondary data collected by the Korea National Statistical Office. Data related to area, employment, housing units, student population, and number of firms were obtained from Korea's Si or Gun's Statistical Year Book 2005. Employment data is divided into basic and non-basic sectors. Basic sectors in Korean economy area farming and manufacturing; and non-basic sector is service sector. Employment in basic sectors allows us to estimate the multiplier effects in the Korean economy. Local revenue and expenditure data were obtained from Korean Local Financial Year Book 2005. Local revenues consist of local tax receipts, other receipts, local share tax, autonomous district control grants, and subsidies (Local finance, MOGAHA). Generally, cities' tax shares are higher (about 50%) than counties' (about 20%). The revenue structure of local governments is presented in Table 1. Population and in-commuters and out-commuters data were obtained from Korean Census of Population data. In this analysis, the cross-sectional data of 172 counties and cities was used for 2005. A list of variables and their summary of statistics such as number of observations, means, standard deviations, and minimum and maximum values for each variable are presented in Table 2.

4. The Spatial Model Estimation Procedure

To begin with, a single equation spatial econometric model is introduced. This is a commonly used model in economics and other fields. Then the features of a simultaneous system of equations are added. The general specification of the single equation is $Y = \lambda WY + X\beta + \varepsilon$

where $\varepsilon = \rho W\varepsilon + u$ (Anselin, 1988). *Y* is a column vector of observations on a dependent variable; *X* is a vector of explanatory variables that are assumed to be uncorrelated with error terms; *W* is a contiguity weight matrix; λ and ρ are spatial lag and spatial error parameters to be estimated; β is the column vector parameters of explanatory variables; and *u* is an independent and identically distributed error term. ε is a spatial error term that can be solved as $\varepsilon = (I - \rho W)^{-1} u$. Spatial dependence has two sources: both error terms and dependent variables may be correlated across space. However, this single equation model does not serve the purpose because a simultaneous system of equation is needed. Following Kelejian and Prucha (2004), a GS3SLS model is specified as follows:

$$Y_n = Y_n B + X_n C + \bar{Y}_n A + U_n \quad (1)$$

$$Y_n = (y_{1,n},\ldots\ldots, y_{m,n}),\ X_n = (x_{1,n},\ldots. x_{k,n}),\ U_n = (u_{1,n},\ldots., u_{m,n}),\ \bar{Y} = (\bar{y}_{1,n},\ldots., \bar{y}_{m,n})$$

$$\bar{y}_{j,n} = W_n y_{j,n} \quad j = 1,\ldots., m \qquad \bar{y}_{ij,n} = \sum_{r=1}^{n} w_{ir,n} y_{rj,n.}$$

where $y_{j,n}$ is the $n \times 1$ vector of cross-sectional observations on the dependent variable, $x_{l,n}$ is the $n \times 1$ vector of cross-sectional observations on the *l*th exogenous variable, $\bar{y}_{j,n}$ is the spatial lag of $y_{j,n}$, $u_{l,n}$ is the $n \times 1$ disturbance vector of in the *j*th equation, W_n is an $n \times n$ weight matrix, and ***B***, ***C***, ***A*** are parameter matrices of mxm, kxm and mxm corresponding variables, respectively. The authors also allow for the spatial autocorrelation in the error term as follows:

$$U_n = \overline{U}_n R + E_n \quad (2)$$

with $E_n = (\varepsilon_{1,n},\ldots,\varepsilon_{m,n})$, $R = diag_{j=1}^{m}(\rho_j)$

$$\overline{U}_n = (\bar{u}_{1,n},\ldots,\bar{u}_{m,n}),\ \bar{u}_{j,n} = W_n u_{j,n} \quad j = 1,\ldots.m,$$

where $\varepsilon_{j,n}$ denotes the column vector of independent and identically distributed error terms and ρ_j denotes the spatial autoregressive parameters. $\bar{u}_{j,n}$ is the spatial lag of $u_{j,n}$.

4.1 GS3SLS Estimation Procedure

The estimation procedure consists of an initial two-stage least squares (2SLS) estimation, followed by estimation of the spatial autoregressive parameter, a generalized spatial 2SLS estimation, and full information estimation (GS3SLS). The first three steps complete the generalized spatial two-stage least squares (GS2SLS), and the final steps take care of the cross-equation error correlation (Kelejian and Prucha, 2004). GS3SLS procedure accounts for two factors. First, it takes care of the simultaneity bias which arises when a dependent variable is correlated with another equation's error term. Second, it allows correlated errors between equations which improved the efficiency of the parameter estimates.

4.2 Initial 2SLS Estimation

The first step in the estimation process consists of the estimation of the model parameter vector β_j in a single-equation spatial econometric model by 2SLS using all exogenous variables, their spatial lag values, and the twice spatially lagged exogenous variables (i.e., X_n, WX_n, W^2X_n). The residual of this step is computed as follows:

$$\tilde{u}_{j,n} = y_{j,n} - Z_{j,n}\tilde{\delta}_{j,n} \tag{3}$$

where $Z_{j,n}$ includes all the endogenous and exogenous variables included in the 2SLS regression.

4.3 Estimation of Spatial Autoregressive Parameter

Equation (2) implies that

$$u_j - \rho_j W u_j = \varepsilon_j \tag{4}$$

and premultiplication of this term by the weights matrix W gives

$$Wu_j - \rho_j W^2 u_j = W\varepsilon_j \tag{5}$$

The following three equation system is obtained from the relationships between equations (4) and (5):

$$\frac{\varepsilon_j^{'}\varepsilon_j}{n} = \frac{u_j^{'}u_j}{n} + \rho_j^2\frac{(Wu_j)^{'}(Wu_j)}{n} - 2\rho_j\frac{u_j^{'}(Wu_j)}{n}$$

$$\frac{(W\varepsilon_j)^{'}(W\varepsilon_j)}{n} = \frac{(Wu_j)^{'}(Wu_j)}{n} - \rho_j^2 \frac{(W^2u_j)^{'}(W^2u_j)}{n} - 2\rho_j \frac{(W^2u_j)^{'}(Wu_j)}{n}$$

$$\frac{\varepsilon_j^{'}W\varepsilon_j}{n} = \frac{u_j(Wu_j)}{n} + \rho_j^2 \frac{(Wu_j)^{'}(W^2u_j)}{n} - \rho_j \frac{u_j^{'}(W^2u_j) + (Wu_j)^{'}(Wu_j)}{n} \tag{6}$$

If the expectations are taken across (6) then the resulting system would be as follows:

$$E\begin{bmatrix} \frac{\varepsilon_j^{'}\varepsilon_j}{n} = \frac{u_j^{'}u_j}{n} + \rho_j^2 \frac{(Wu_j)^{'}(Wu_j)}{n} - 2\rho_j \frac{u_j^{'}(Wu_j)}{n} \\ \frac{(W\varepsilon_j)^{'}(W\varepsilon_j)}{n} = \frac{(Wu_j)^{'}(Wu_j)}{n} - \rho_j^2 \frac{(W^2u_j)^{'}(W^2u_j)}{n} - 2\rho_j \frac{(W^2u_j)^{'}(Wu_j)}{n} \\ \frac{\varepsilon_j^{'}W\varepsilon_j}{n} = \frac{u_j(Wu_j)}{n} + \rho_j^2 \frac{(Wu_j)^{'}(W^2u_j)}{n} - \rho_j \frac{u_j^{'}(W^2u_j) + (Wu_j)^{'}(Wu_j)}{n} \\ \end{bmatrix} \tag{7}$$

$$E\begin{bmatrix} \sigma_{\mathrm{j}}^2 \\ \rho_{\mathrm{j}}^2 \frac{tr(W'W)}{n} \\ \rho_{\mathrm{j}}^2 \frac{tr(W)}{n} = 0 \end{bmatrix} = E\begin{bmatrix} \frac{u_j^{'}u_j}{n} + \rho_j^2 \frac{(Wu_j)^{'}(Wu_j)}{n} - 2\rho_j \frac{u_j^{'}(Wu_j)}{n} \\ \frac{(Wu_j)^{'}(Wu_j)}{n} - \rho_j^2 \frac{(W^2u_j)^{'}(W^2u_j)}{n} - 2\rho_j \frac{(W^2u_j)^{'}(Wu_j)}{n} \\ \left[\frac{u_j(Wu_j)}{n} + \rho_j^2 \frac{(Wu_j)^{'}(W^2u_j)}{n} - \rho_j \frac{u_j^{'}(W^2u_j) + (Wu_j)^{'}(Wu_j)}{n}\right] \end{bmatrix} \tag{8}$$

$$E\begin{bmatrix} \frac{u_j^{'}u_j}{n} \\ \frac{(Wu_j)^{'}(Wu_j)}{n} \\ \frac{u_j(Wu_j)}{n} \end{bmatrix} = E\begin{bmatrix} 2\rho_j \frac{E(u_j^{'}(Wu_j))}{n} & -\rho_j^2 \frac{E((Wu_j)^{'}(Wu_j))}{n} & \sigma_{\mathrm{j}}^2 \\ 2\rho_j \frac{E((W^2u_j)^{'}(Wu_j))}{n} & -\rho_j^2 \frac{(W^2u_j)^{'}(W^2u_j)}{n} & \rho_{\mathrm{j}}^2 \frac{tr(W'W)}{n} \\ \rho_j \frac{E(u_j^{'}(W^2u_j) + (Wu_j)^{'}(Wu_j))}{n} & -\rho_j^2 \frac{E((Wu_j)^{'}(W^2u_j))}{n} & 0 \end{bmatrix}$$

(9)

$$E\begin{bmatrix} \frac{u_j^{'}u_j}{n} \\ \frac{(Wu_j)^{'}(Wu_j)}{n} \\ \frac{u_j(Wu_j)}{n} \end{bmatrix} = E\begin{bmatrix} 2\frac{E(u_j^{'}(Wu_j))}{n} & -\frac{E((Wu_j)^{'}(Wu_j))}{n} & 1 \\ 2\frac{E((W^2u_j)^{'}(Wu_j))}{n} & -\frac{(W^2u_j)^{'}(W^2u_j)}{n} & \frac{tr(W'W)}{n} \\ \frac{E(u_j^{'}(W^2u_j)+(Wu_j)^{'}(Wu_j))}{n} & -\frac{E((Wu_j)^{'}(W^2u_j))}{n} & 0 \end{bmatrix}\begin{bmatrix} \rho_j \\ \rho_j^2 \\ \sigma_j^2 \end{bmatrix} \tag{10}$$

The right-hand side of equation (10) can be written in the following form:

$$\frac{1}{n}\begin{bmatrix} 2\tilde{u}'_{j,n}\tilde{\tilde{u}}_{j,n} & -\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{u}}_{j,n} & n \\ 2\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & -\tilde{\tilde{\tilde{u}}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & Tr(W_n'W_n) \\ \tilde{u}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n}+\tilde{\tilde{\tilde{u}}}'_{j,n}\tilde{\tilde{u}}_{j,n} & -\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & 0 \end{bmatrix}\begin{bmatrix} \rho_j \\ \rho_j^2 \\ \sigma_j^2 \end{bmatrix} \tag{11}$$

This system of equation can be written as

$$\gamma_{j,n} = \Gamma_{j,n}\alpha_j \quad \rightarrow \alpha_j = \Gamma_{j,n}^{-1}\gamma_{j,n} \tag{12}$$

where $\Gamma_{j,n} = \frac{1}{n}\begin{bmatrix} 2\tilde{u}'_{j,n}\tilde{\tilde{u}}_{j,n} & -\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{u}}_{j,n} & n \\ 2\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & -\tilde{\tilde{\tilde{u}}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & Tr(W_n'W_n) \\ \tilde{u}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n}+\tilde{\tilde{\tilde{u}}}'_{j,n}\tilde{\tilde{u}}_{j,n} & -\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & 0 \end{bmatrix}$

The parameter vector $\rho_j, \rho_j^2, \sigma_j^2$ would be determined in terms of the relation in equation (12).

They minimize the following equation:

$$\left[g_{j,n} - \Gamma_{j,n}\begin{bmatrix} \rho_{j,n} \\ \rho_{j,n}^2 \\ \sigma_{jj} \end{bmatrix}\right]'\left[g_{j,n} - \Gamma_{j,n}\begin{bmatrix} \rho_{j,n} \\ \rho_{j,n}^2 \\ \sigma_{jj} \end{bmatrix}\right]$$

$$\Gamma_{j,n} = \frac{1}{n}\begin{bmatrix} 2\tilde{u}'_{j,n}\tilde{\tilde{u}}_{j,n} & -\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{u}}_{j,n} & n \\ 2\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & -\tilde{\tilde{\tilde{u}}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & Tr(W_n'W_n) \\ \tilde{u}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n}+\tilde{\tilde{\tilde{u}}}'_{j,n}\tilde{\tilde{u}}_{j,n} & -\tilde{\tilde{u}}'_{j,n}\tilde{\tilde{\tilde{u}}}_{j,n} & 0 \end{bmatrix}, \qquad g_{j,n} = \frac{1}{n}\begin{bmatrix} \tilde{u}'_{j,n}\tilde{u}_{j,n} \\ \tilde{\tilde{u}}'_{j,n}\tilde{\tilde{u}}_{j,n} \\ \tilde{u}'_{j,n}\tilde{\tilde{u}}_{j,n} \end{bmatrix} \tag{13}$$

Where $\tilde{\tilde{u}}_{j,n} = W_n\tilde{u}_{j,n}$ and $\tilde{\tilde{\tilde{u}}}_{j,n} = W_n^2$.

4.4 Estimation of GS2SLS

In this stage, a Cochrane-Orcutt type transformation is applied to dependent, endogenous, and exogenous variables of the single-equation spatial econometric model by using estimated spatial autoregressive parameters to account for the spatial correlation. Let $y^*_{j,n} = y_{j,n} - \tilde{\rho}_j W_n y_{j,n}$ and $Z^*_{j,n} = Z_{j,n} - \tilde{\rho}_j W_n Z_{j,n}$. Then the equation becomes:

$$y^*_{j,n}(\rho_j) = Z^*_{j,n}(\rho_j)\delta_j + \varepsilon_{j,n} \quad \rightarrow \quad \hat{\delta}_{j,n} = \left[\hat{Z}^*_{j,n}(\rho_j)'\hat{Z}^*_{j,n}(\rho_j)\right]^{-1}\hat{Z}^*_{j,n}(\rho_j)' y^*_{j,n}(\rho_j) \qquad (14)$$

where $\hat{Z}^*_{j,n}(\rho_j) = P_H Z^*_{j,n}(\rho_j)$ with $P_H = H_n(H'_n H_n)^{-1}H'_n$ assuming ρ_j is known. This $\hat{\delta}_{j,n}$ becomes the GS2SLS estimator. The feasible GS2SLS estimator for ρ_j, (say $\hat{\rho}^F_j$) is now defined by substituting the generalized moments estimator $\tilde{\rho}_{j,n}$ for ρ_j in equation (14), that is

$$\hat{\delta}^F_{j,n} = \left[\hat{Z}^*_{j,n}(\tilde{\rho}_j)'\hat{Z}^*_{j,n}(\tilde{\rho}_j)\right]^{-1}\hat{Z}^*_{j,n}(\tilde{\rho}_j)' y^*_{j,n}(\tilde{\rho}_j) \qquad (15)$$

4.5 Full Information Estimation (GS3SLS)

Up to this point, our model accounts for the potential spatial correlation, but it does not take into account the potential cross-equation correlation in the innovation vector ε_j. To account for this, it is helpful to stack the equations in (14) as

$$y^*_n(\rho) = Z^*_n(\rho)\delta + \varepsilon_n \qquad (16)$$

where

$$y^*_n(\rho) = (y^*_{1,n}(\rho_1)',....,y^*_{m,n}(\rho_m)')',$$

$$Z^*_n(\rho) = diag^m_{j=1}(Z^*_{j,n}(\rho_j)) \;\; ; \quad \rho = (\rho_1,....,\rho_m)' \text{ and } \delta = (\delta'_1,.....,\delta'_m)'$$

It is assumed that $E\varepsilon_n = 0$ and $E\varepsilon_n\varepsilon'_n = \Sigma \otimes I_n$. If ρ and Σ are known, a natural system of instrumental variables estimator of δ would be

$$\breve{\delta}_n = [\hat{Z}_n^*(\rho)'(\Sigma^{-1} \otimes I_n)(\hat{Z}_n^*(\rho)]^{-1} \hat{Z}_n^*(\rho)'(\Sigma^{-1} \otimes I_n) y_n^*(\rho) \quad (17)$$

Where $\hat{Z}_n^*(\rho) = diag_{j=1}^m(\hat{Z}_{j,n}^*(\rho_j))$ and $\hat{Z}_{j,n}^*(\rho_j) = P_H Z_{j,n}^*(\rho_j)$

To estimate equation (16), the estimators for ρ and Σ are needed to be found. Let

$\tilde{\rho}_n = (\tilde{\rho}_{1,n}, ..., \tilde{\rho}_{m,n})'$ where $\tilde{\rho}_{j,n}$ denotes the generalized moments estimator for ρ_j. The

consistent estimator of Σ is $\hat{\Sigma}_n$ where $\hat{\Sigma}_n$ is estimated as a $m \times m$ matrix whose $(j, l)^{\text{th}}$ element

is $\hat{\sigma}_{jl,n} = n^{-1} \tilde{\varepsilon}_{j,n} \tilde{\varepsilon}_{l,n}$ and $\tilde{\varepsilon}_{j,n} = y_{j,n}^*(\tilde{\rho}_{j,n}) - Z_{j,n}^*(\tilde{\rho}_{j,n}) \hat{\delta}_{j,n}^F$. (18)

Replacing the value consistent estimator in equation (17), a feasible GS3SLS estimator is obtained as

$$\breve{\delta}_n^F = [\hat{Z}_n^*(\tilde{\rho})'(\hat{\Sigma}^{-1} \otimes I_n)(\hat{Z}_n^*(\tilde{\rho})]^{-1} \hat{Z}_n^*(\tilde{\rho})'(\hat{\Sigma}^{-1} \otimes I_n) y_n^*(\tilde{\rho}_n) \quad (19)$$

5. Model Specification

The model developed, specified, and estimated below is a combination of labor markets, housing markets, demography, and local public finance variables; however, the labor markets play the central role in this modeling framework. The model is built on the assumption that economic growth is largely caused by an exogenous increase in employment. Employers create local jobs while the residential choices of employees create local labor forces. Each employer faces a short-run labor supply within commuting distance from the plant, known as the commuting shed. Other employers within the commuting shed share the same workforce. Similarly, each member of the local labor force faces a demand for labor that consists of the sum of all jobs within his or her commuting shed. Also, other workers within the commuting shed share the same labor demand forces but may be subject to labor demands form outside the commuting shed of the first worker.

Individual workers make residential decisions based on job availability, relative costs of living, local amenities, quality of public services, and other items that affect their quality of life. The workers also choose among available jobs based on skill requirements, wage rates, job security, and commuting costs. As the same time, employers locate their plants based on cost of doing business, marketing considerations, and the availability of workers and other resources. The labor market allocates jobs among the currently employed, in-commuters, out-commuters, and in-migrants. Some new jobs are also taken by currently employed workers who change positions.

As Tiebout (1956) points out, the workers also choose a residence community that offers a mix of local public goods and services best suited to their tastes. By choosing to relocate, or "voting with their feet," consumers reveal their preferences for local public goods. Together with the labor market and public goods market equilibrium, the population of local areas is determined.

Our model has 10 structural equations, each of which has the following general form that is similar to the standard Cliff and Ord (1973) type model with a spatial dependent variable lag and a spatial autoregressive error term:

$$\begin{aligned} Y_k &= a_{k,0} + \lambda_k \overline{Y}_k^N + \sum_{j=1}^{J} \beta_{k,j} Z_{k,j} + \varepsilon_k \\ \varepsilon_k &= \rho_k W \varepsilon_k + u_k \end{aligned} \tag{20}$$

For the k^{th} equation, Y_k is an $n \times 1$ column vector of observations on the dependent variable, $\overline{Y}_k^N = W \times Y_k$ is an $n \times 1$ column vector of observations on the spatially weighted averages of the dependent variable, $Z_{k,j}$ is an $n \times j$ matrix of observations on the endogenous and exogenous variables for the k^{th} dependent variable, W is an $n \times n$ matrix of spatial weights that relate all locations in our cross-section sample to their neighboring locations. The parameters a, λ, β, and ρ are the GS3SLS estimates, ε is a spatially related regression error term, and u is a regression error term with the usual independent and identically distributed statistical properties. Furthermore, the spatially related error term (ε_k) can be solved in terms of ρ_k and W:

$\varepsilon_k = (I - \rho_k W)^{-1} u_k$.

The spatial lag variables for a county are defined as the weighted average values for the set of neighboring counties (i.e., if they are located with 30 km of radial distance). These neighbors' average values for all geographic units are computed by post-multiplying a "row-normalized" spatial weights matrix by a column vector of cross-sectional observations of a variable. A spatial weights matrix is a square matrix that relates each cross-sectional unit to its unique set of neighbor areas. A row-normalized spatial weights matrix (W) is one whose row sums are all equal to one. Three types of row-normalized spatial weights matrices are investigated in this paper.

First, a "simple" gravity (weighted inverse distance) row-normalized spatial weights matrix is used, whose typical values are

$$W_{G1} = [w_{ij}] = \frac{X_j / D_{ij}}{\sum_{J=1}^{N} X_J / D_{iJ}} \tag{21}$$

if j is a neighbor of i, otherwise $w_{ij} = 0$.

Second, a more typical gravity row-normalized (weighted inverse distance squared) spatial weights matrix was used, whose common elements are

$$W_{G2} = [w_{ij}] = \frac{X_j / D_{ij}^2}{\sum_{J=1}^{N} X_J / D_{iJ}^2} \tag{22}$$

if j is a neighbor of i, otherwise $w_{ij} = 0$.

The weight variable in the gravity calculation (X) is used to account for issues related to size or mass. Larger and heavier objects are more attractive than are smaller and lighter objects, and places closer together have greater attraction. To measure this size or mass, employment total is used. D_{ij} is the distance (in miles) between locations i and j. Distance is calculated from one population centroid to another.

Third, a row-normalized spatial weights matrix with uniform values was used, whose typical values are $W_U = [w_{ij}] = 1/N_i$ if location j is within 30 km from location i or $w_{ij} = 0$ if not (N_i is the number of location i 's neighbors).

As Kelejian and Robinson (1995) point out, two of the estimated parameters in equation (20) need special mention: λ_k and ρ_k (the spatial lag and spatial autoregressive parameters, respectively). The parameter spaces for these estimated coefficients have a restricted range: $1/\psi^*_{k,neg} < \lambda_k, \rho_k < 1/\psi^*_{k,pos}$, where $\psi^*_{k,neg}$ is the largest negative eigenvalue of the spatial weights matrix (W) and $\psi^*_{k,pos}$ is the smallest positive eigenvalue of the spatial weights matrix (W). This range will always provide a "clear" parameter space that includes the value zero. If the spatial weights matrix is "row-normalized" (i.e., the row elements of W sum to 1 or, in other words, form a proportion distribution) then the smallest positive eigenvalue will always equal 1 ($\psi^*_{k,pos} = 1$). However, except for a few theoretical types of spatial weights matrices, the largest negative eigenvalue is greater than −1 ($-1 < \psi^*_{k,neg} < 0$). This means that the parameter space for the spatial lag and spatial autoregressive parameters will be, in general, between some value less than −1 and +1 when the spatial weights matrix is row-normalized.

The expanded version of equation (20) is as follows. The expected signs are presented in the parenthesis just below each explanatory variable.

$$REVLOC = \alpha_1 + \alpha_2 WREVLOC + \alpha_3 POPTOT + \alpha_3 EMPNBAS + \alpha_4 COMIN + \varepsilon$$

(−) (+) (+) (−)

$$EXPLOC = \quad \beta_1 + \beta_2 WEXPLOC + \beta_3 POPTOT + \beta_4 COMIN + \beta_5 EMPNBAS + \varepsilon$$

(±) (+) (+) (+)

$$HOUSTOT = \delta_1 + \delta_2 WHOUSTOT + \delta_3 POPTOT + \delta_4 COMIN + \varepsilon$$

(±) (+) (−)

$$POPTOT = \gamma_1 + \gamma_2 WPOPTOT + \gamma_3 POPEAP + \varepsilon$$

(±) (+)

$$POPEAP = \lambda_1 + \lambda_2 WPOPEAP + \lambda_3 POPTOT + \lambda_4 EMPNBAS + \varepsilon$$

(±) (+) (+)

$$STDTTOT = \mu_1 + \mu_2 WSTDTTOT + \mu_3 POPTOT + \varepsilon$$

(±) (+)

$$COMOUT = \nu_1 + \nu_2 WCOMOUT + \nu_3 POPEAP + \nu_4 EMPTOT + \nu_5 AREA + \nu_6 AEMP +$$

(±) (+) (−) (+) (±)

$$\nu_7 CEMP + \nu_8 EXPED + \varepsilon$$

(+) (−)

$$COMIN = \pi_1 + \pi_2 WCOMIN + \pi_3 POPEAP + \pi_4 EMPTOT + \pi_5 CEMP + \pi_6 AREA + \pi_7 AEMP + \varepsilon$$

(±) (−) (+) (−) (+) (±)

$$FIRMTOT = \rho_1 + \rho_2 WFIRMTOT + \rho_3 POPTOT + \rho_4 AREA + \rho_5 AEMP + \varepsilon$$

(±) (+) (+) (±)

$$EMPNBAS = \sigma_1 + \sigma_2 WEMPNBAS + \sigma_3 EMPTOT + \sigma_4 AREA + \sigma_5 EXPED + \sigma_6 AEXPED +$$

(±) (+) (+) (+) (±)

$$\sigma_7 CEXPED + \varepsilon$$

(+)

Variables' names and descriptions are presented in Table 2.

6. Model Estimation Results and Discussion

As mentioned in the previous section, a GS3SLS procedures is applied to estimate the parameter value of our model, which consists of 10 spatially interrelated simultaneous equations.

Before estimating the model using GS3SLS, the different models were estimated using ordinary least squares (OLS) and GS2SLS and evaluated individually based on several criteria (adjusted R^2, correct signs, statistical significance). Table 3 presents the parameters estimates and their significance by OLS, GS2SLS, and GS3SLS procedures. Overall the signs of the parameter estimates appeared to be robust; however the magnitudes of these estimates vary across different estimation procedure used. In some cases, not only the magnitude of the coefficient change, but also the sign of the coefficients flip as the estimation procedure is changed. For example, the variable in-commuters appears to be significant and positive in explaining local revenue in OLS model, however it is negative and significant in GS3SLS model. The same variable appears to be non-significant in local expenditure equation using OLS procedure, however it is highly significant and negative in GS3SLS. Likewise, spatial lag of local revenue is negative and significant in local revenue equation when used OLS procedure, however it is positive and significant when used GS3SLS procedure. This shows that when spatial interaction and cross-equation interaction are taken into account, the unbiased parameter estimates may be estimated. Compared to GS2SLS, it is found that the magnitude of the GS3SLS coefficients of many explanatory variables appear to have changed significantly. Most estimated parameter values are significant at the 1% level and a few at the 5% and 10% levels. The fact that all equations have a R^2 value higher than 0.90 indicates that our model possesses reasonably high explanatory power. Most of the estimated parameter values have signs that would be expected or can be explained. The results are based on a gravity-based row-normalized weight matrix (weight divided by distance); however, different weight matrices have been tried. A separate section is devoted to a detailed sensitivity analysis of weight matrices. The following results and their interpretations are based on simultaneous spatial lag model (SL3SLS).

The majority of spatial lags of dependent variables appear to be significant at the 5% level. This shows evidence of significant spatial spillover in Korean regions. Based on the magnitude of the estimated coefficients, the positive spatial pattern appears to be strongest for population, students' numbers, out-commuters, and in-commuters. The negative spatial pattern appears to be strongest for local expenditures and number of firms. The negative sign of the spatial lag of local expenditure supports previous studies in which local public expenditures (e.g., transport, education, parks and recreation) have spillover effects in the neighborhood (Case, Hines, and Resen, 1993; Murdoch, Rahmatian, and Thayer, 1993).

The spatial lag of local revenue is insignificant. This may be partly due to (1) overdependence of local regions on central government for revenue generation; (2) less flexibility on the part of local government for policy making; and (3) formula-driven revenue collection. The coefficients of spatial lag of population and the number of students are not significant. As expected, spatial dependence for in-commuting and out-commuting is positive and significant.

As expected, local revenue is positively and significantly impacted by population, and non-basic employment, whereas it is negatively impacted by in-commuting. Local expenditure is found to be dependent on and influenced by population and employment in non-basic sectors significantly and positively. Unexpectedly, it is negatively impacted by in-commuting. It is hypothesized that in-commuting will exert a positive influence on local expenditure because it represents the daytime population of the area, and higher demand for services by employer which further pushes the demand for local services. This may be due to lack of an income variable in the local expenditure equation. Due to the lack of an income variable, it is not possible to test the hypothesis that more affluent communities demand higher quality services and are more willing to pay for them.

The total housing units equation is estimated as a function of spatial lag of itself, population, and in-commuting. As expected, total population is found to be the most important determinant of housing units. It is estimated that in-commuting is significantly and negatively associated with total housing units; which is expected.

A population equation is estimated as a function of spatial lag of itself and economically active population. Because labor force data is not available, the economically active population is used as a proxy of labor force. As expected, economically active population is found to be significant to explain the population.

As a proxy of labor force, economically active population equation is estimated as a function of spatial lag of itself, population, and employment in non-basic sectors. As expected, economically active population is positively impacted by population and employment in non-basic sectors. Student numbers is estimated as a function of spatial lag of itself and population. As expected, student number is positively and significantly impacted by population.

Out-commuting equation is estimated as a function of spatial lag of itself, economically active population, employment, area, area×employment, external employment, and economic development expenditures. The variable area × employment is also called as expansion variable which captures the structural changes that are caused by the different sizes of counties. All variables are significant. The signs of all variables are as expected. It appears that external employment drives the out-commuting up whereas economic development expenditures drive it down. The estimated parameter of the expansion variable (area × employment) is significant and negative. This implies that for larger counties, the area variable decreases the marginal effect of employment on out-commuting even though the area variable alone may not be significant. In this case, the area variable is also significant and positive. As expected, the positive sign of the coefficient for the area variable shows that as the county area increases, out-commuting also increases.

In-commuting is estimated as a function of the spatial lag of itself, economically active population, employment, external employment, area, area × employment. The signs of all variables are as expected except for external employment. An increase in economically active population tends to decrease the in-commuting; which is logical. As expected, the employment variable is found to be significant and positive. This implies that increased employment opportunities in residence counties create increased in-commuting. Surprisingly, the external employment is positive and significant. Area variable has an unexpected positive sign but it not significant. The expansion variables (area × employment) have a significant and negative coefficient. It appears that for larger counties the area variable decreases the marginal effect of employment on in-commuting.

The number of firms is modeled as a function of spatial lag of itself, population, area, and area×employment. All variables have expected signs and are significant. It appears that number of firms increases as the area, population, area × employment increase. As expected, the number of firm variable is strongly and negatively impacted by spatial lag of itself. This implies that there is competition among firms located in residence counties and neighboring counties.

Employment in non-basic sectors is estimated as a function of the spatial lag of itself, employment total, area, economic development expenditures, external economic development expenditures, and expansion variable—area × economic development expenditures. All explanatory variables are found to be significant except spatial lag of dependent variable and external economic development expenditures. The negative sign of the spatial lag variable indicates that increased non-basic employment in neighboring regions negatively impacts the residence county. As expected, employment total variable significantly and positively impacts non-basic employment. The economic development expenditures appear to impact significantly and positively. The coefficient of area × economic development expenditures appear to be negative and significant. This implies that for larger counties the area variable decreases the marginal effect of economic development expenditure on non-basic employment.

6.1 Sensitivity of Choice of Spatial Linkages

The many alternative methods of specification of spatial linkages creates difficulties and controversies in spatial data analysis. Sensitivity analysis is used to determine how "sensitive" a model is to changes in the weight matrices representing different spatial linkages. It is possible to build the confidence in the model by studying uncertainties that are associated with different weight matrices. As mentioned earlier, three spatial weight matrices are used based on distance, weight and inverse distance, and weight and inverse distance squared. The latter two are also called gravity based matrices. The total employment is used as a weight variable. Using different matrices, the model (Table 4) is estimated using GS3SLS procedure. Overall results appear to be robust across different spatial linkages**.** However, the magnitude and significance of coefficients of some variables are found to be sensitive to the choice of spatial matrices. For example, spatial lag of local expenditure is significant and negative in the model that used the gravity based weight matrices, whereas it is not significant in the model that used uniform weight matrix. The spatial lag of economically active population is negative and significant when used with uniform weight matrices, however it is not significant when gravity based weight matrices were used. However the sign remains the same. Likewise, spatial lag of employment in non-basic sectors is significant and negative in a model that used uniform spatial weight matrix whereas the same variable is not significant in both the models that used gravity based weight matrices. In some cases, it appears that the magnitude of the variables are also changed, which shows that analytical results may be sensitive to the specification of spatial weight matrix.

6.2 Model Validation

The fact is that in general model validation and model building processes move together. Before deciding on a "ideal" model, MAPE was used as a measure of the forecast accuracy to evaluate these models (Table 7). Based on MAPE criterion used in in-sample data, it appears that not all equations consistently perform well (see Table 8). Predictive accuracy of local revenue, local expenditures, population, out-commuters, number of firms and employment in non-basic sectors are found to be better in the SLE3SLS model, whereas housing units, economically active population, total students are better forecasted by the SL3SLS model. None of the equations in a SE3SLS model and NS3SLS model have better forecasting accuracy than the SLE3SLS model, and SL3SLS model except in-commuting variable. This implies that there exists a significant spatial spillover effect in Korean local economies. Although both SL3SLS and SLE3SLS models appeared to be similar in terms of overall MAPE statistic, SL3SLS model have an advantage of being parsimonious. Another advantage of SL3SLS over SLE3SLS is that the reduced form solutions are easy to handle and make intuitive sense. Therefore, reduced form estimates is estimated using structural equation obtained from SL3SLS model.

6.2 Reduced Form Estimates

The reduced form equations are obtained by solving structural equations derived from SL3SLS model. In this case, all endogenous variables are functions of exogenous variables. Solving spatial structural equations to obtain a reduced form equation is a daunting task.

However, by following Kelejian and Prucha (2004), we[1] obtained a reduced form estimate of spatial simultaneous lag model which is as follows.

$$Y_n = Y_n B + X_n C + \bar{Y}_n A + U_n \tag{1}$$

$$Y_n = (y_{1,n},\ldots\ldots,y_{m,n}),\; X_n = (x_{1,n},\ldots\ldots x_{k,n}),\; U_n = (u_{1,n},\ldots\ldots,u_{m,n}),\; \bar{Y} = (\bar{y}_{1,n},\ldots\ldots,\bar{y}_{m,n})$$

$$\bar{y}_{j,n} = W_n y_{j,n} \quad j = 1,\ldots\ldots,m \qquad \bar{y}_{ij,n} = \sum_{r=1}^{n} w_{ir,n} y_{rj,n}.$$

where *n* is regions, *m* endogenous variables, and *k* exogenous variable. The dimension of coefficient as follows: $B_{(m\times m)}$, $C_{(k\times m)}$, $A_{(m\times m)}$

$$y_n = vec(Y_n) = \begin{pmatrix} y_{1,n} \\ \vdots \\ \vdots \\ y_{m,n} \end{pmatrix} \qquad x_n = vec(X_n) = \begin{pmatrix} x_{1,n} \\ \vdots \\ \vdots \\ x_{k,n} \end{pmatrix}$$

If A_1 and A_2 are conformable matrices, then $vec(A_1 A_2) = (A_2^{'} \otimes I)vec(A_1)$ (Berck et al 1993).

Following this rule, reduced form solution of equation (1) would be as follows.

$$y_n = \left[(B'\otimes I_n) + (A'\otimes W_n)\right]y_n + (C'\otimes I_n)x_n \tag{23}$$

$$y_n = \left\{(I_m \otimes I_n) - [(B'\otimes I_n) + (A'\otimes W_n)]\right\}^{-1}(C'\otimes I_n)x_n \tag{24}$$

The dimension of coefficient matrices are as follows:

$$B'_{(m\times m)} = \begin{bmatrix} 0 & \beta_{12} & \cdots\cdots & \beta_{1m} \\ \beta_{21} & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ \vdots & & \ddots & \vdots \\ \beta_{m1} & \cdots & \beta_{m,m-1} & 0 \end{bmatrix} \qquad A'_{m\times m} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots\ 0 & \lambda_m \end{bmatrix}$$ when only the dependent variable has spatial lag.

Reduced form estimate is very huge to present in a table, however the results of impact estimates are presented in the next section.

[1] Dr. Dennis Robinson with the help of Kelejian and Prucha (2004) developed the equation (24).

6.3 Impact Estimation

The uniqueness of the solution is that cross-county spatial spillover effect can be estimated through this model. Once the reduced form equation is obtained, it can be used for impact analysis purpose. Although there are seven exogenous variables in the model, employment is one of the main driving forces for all sectors of the economy. Changes in employment lead to increases in population and wage levels, which ultimately alter demands for public services and the revenues available to fund these services. To demonstrate how the model works, and to determine a reasonable estimate of the impact, a 1000 jobs increase is hypothesized in Gujang county of Punsan Province. As a test community, Gijang County is a fairly small county with a population of approximately 73,000. Using the reduced form coefficients of our model, the impacts of an employment change was estimated on the Gijang County economy. Change in employment also changes the external employment for other counties. It also affects the expansion variable area × employment. After accounting for these changes, the impact of 1,000 new jobs was estimated using a reduced form equation of spatial lag simultaneous equation model.

The creation of additional jobs caused an increase in local revenue of 3.6 billion won and an increase in expenditures of 3.2 billion won. The new jobs also caused the following increases:

- total housing units by 482;
- economically active population by 1660;
- number of students by 356;
- in-commuters by 319;
- out-commuters by 267;
- number of firms by 191;

- non-basic employment by 813.

Also estimated was the spatial spillover effect of additional 1000 jobs to neighboring regions. It is found that 10 neighboring counties impacted by this change in employment; however only four counties appeared to have sizable impact (Table 6). Once the spatial impact spillover to counties (other than Gijang) is removed, the impacts estimated from spatial and non-spatial model can be compared. The results show that the intra-county impact estimates of the non-spatial model are 10% to 9% lower than the impacts estimated by the spatial model for four dependent variables (local revenue, local expenditure, housing units, and number of firms) (Table 6b). This implies that if the spatial spillover effect is ignored, we may be underestimating the impact on these variables. In in-and out-commuting, non-spatial model overestimate the impact- 19% and 24% respectively. Population related variables such as total population, total students and economically active population, both model predictions found to be comparable. This implies that there is little spatial spillover effect in these variables. This shows that accounting for spatial interactions is imperative to improve model performance.

7. Summary and Conclusion

As stated previously in the research objective, the primary goal of this study was to develop and estimate a model that accounts for the cross-county and cross-equation spillover effects in local regions in Korea. The different versions of a spatial model is estimated, that accounts for the interregional spillover effect. The model-building process began with the estimation and evaluation of each equation using criteria such as R^2 and *t*-statistics. Then all equations were collapsed into one system and estimated model using GS3SLS. Before finalizing, the model was validated based on predictive accuracy as measured by MAPE statistic. A SL3SLS model is found to be the ‘best’ model for Korean regions. The model contains equations for local revenue, local expenditures, total housing units, population total, economically active population, number of students l, in- and out-commuting, total firms, and non-basic employment and assumes that employment is the main driver of the Korean regional economy. Other exogenous variables included in this model were economic development expenditures, area, and expansion variables. It is found that a significant cross-county and cross-equation spillover effects exist in Korean regions. The results in some cases appear to be sensitive to the choice of spatial linkages as defined by weight matrices.

Table 1. Revenue Structure of Local Government in Korea

	Year				
Sources of Revenue (%)	2000	2001	2002	2003	2004
Local tax	35.3	37.2	37.4	34.3	35.1
Non-tax revenue	28.9	25.6	26.2	33.1	34.5
Transfer revenue from central government	34.4	35.8	35.6	32.1	29.8
Local bond	1.4	1.1	0.7	0.6	0.6

Sources: *Financial Yearbook of Local Government, 2000-2004*.

Table 2. Variables, Variable Descriptions, and Descriptive Statistics[a]

Variable	Label	Mean	SD	Minimum	Maximum
AREA	Area in square kilometers	591	334	33	1818
POP_TOT2	Population squared (million)	792206	7429204	69	96435758
EMP_TOT	Total employed people	118789	357844	4871	4363664
A_EMP	Area × employment	71211531	230541742	353440	2640000000
EXP_ED	Economic development expenditures	126575	256145	16683	2874672
C_EMP	External employment	676840	1461423	0	7237904
A_EXPED	Area × economic development expenditures	80169595	177829596	1012324	1740000000
REV_LOC	Local Annual Revenue	638431	1480464	51330	17800000
EXP_LOC	Local annual expenditures	495175	1342015	91257	16200000
HOUS_TOT	Total housing units	75301	206345	2890	2321949
POP_TOT	Total population	277547	848148	8331	9820171
POP_EAP	Economically active population	197650	637780	5597	7432406
STDT_TOT	No. of total students	48960	138099	1193	1528649
COM_OUT	Out-commuters	20950	50672	0	464489
COM_IN	In-commuters	21122	79203	0	1006101
FIRM_TOT	Total number of firms	31010	79595	1914	946620
EMP_NBAS	Non basic employment other than farm and manufacturing employment	77538	313949	3119	3878251
W_REV_LOC	Spatial lag of local revenue	671322	662841	0	3336419
W_EXP_LOC	Spatial lag of local expenditures	528140	572401	0	2902942
W_HOUS_TOT	Spatial lag of total housing units	80325	97291	0	470909
W_POP_TOT	Spatial lag of total population	301585	398613	0	1890502
W_POP_EAP	Spatial lag of economically active population	216128	297233	0	1403936
W_STDT_TOT	Spatial lag of total students	52591	67634	0	317318
W_COM_OUT	Spatial lag of out-commuters	22754	36490	0	153646
W_COM_IN	Spatial lag of in-commuters	24781	41251	0	180649
W_FIRM_TOT	Spatial lag of firm total	33087	33851	0	169629
W_EMP_NBAS	Spatial lag of employment in non-basic sectors	87971	132658	0	630065

[a] $N = 172$

Table 3. Regression Results: Ordinary Least Squares and Generalized Spatial Two-Stage and Three-Stage Least Squares

Model	Variables	OLS Estimates	*p*-value	GS2SLS Estimates	*p*-value	GS3SLS Estimates	*p*-value
Local Revenue	Intercept	208493.2	<.0001	206064.8	<.0001	215844.7	<.0001
	W_REV_LOC[a]	-0.01289	0.0004	-0.00799	0.0619	0.000245	0.9495
	POP_TOT	1.21358	<.0001	1.190056	<.0001	0.9272	<.0001
	EMP_NBAS	1.206571	<.0001	1.645595	0.0003	2.797318	<.0001
	COM_IN[a]	0.984833	0.0241	-0.58531	0.3378	-2.46465	<.0001
	Adj R^2	0.99311		0.99246			
Local Expenditure	Intercept	133054.6	<.0001	122873.5	<.0001	125231	<.0001
	W_EXP_LOC	-0.01538	<.0001	-0.01398	0.0002	-0.00759	0.0199
	POP_TOT	0.77954	<.0001	0.93104	<.0001	0.828126	<.0001
	COM_IN[a]	-0.35441	0.2903	-1.27861	0.0087	-2.75825	<.0001
	EMP_NBAS	2.256863	<.0001	2.072154	<.0001	2.696653	<.0001
	Adj R^2	0.99496		0.9942			
Total Housing Units	Intercept	6814.276	<.0001	6323.118	<.0001	6234.474	<.0001
	W_HOUS_TOT	0.001455	0.4146	0.002575	0.1666	0.002704	0.1291
	POP_TOT	0.269944	<.0001	0.276472	<.0001	0.278249	<.0001
	COM_IN	-0.31962	<.0001	-0.39366	<.0001	-0.41414	<.0001
	Adj R^2	0.99794		0.99779			
Population	Intercept	13715.63	<.0001	13833.41	<.0001	13857.61	<.0001
	W_POP_TOT	0.001232	0.0646	0.001251	0.0609	0.00118	0.0749
	POP_EAP	1.329447	<.0001	1.328766	<.0001	1.328956	<.0001
	Adj R^2	0.99969		0.99969			
Economically Active Population	Intercept	-6524.46	<.0001	-7211.04	<.0001	-8084.37	<.0001
	W_POP_EAP	0.000526	0.1009	0.000198	0.5647	-0.00027	0.4188
	POP_TOT	0.678005	<.0001	0.69121	<.0001	0.70867	<.0001
	EMP_NBAS	0.201956	<.0001	0.166249	<.0001	0.118865	<.0001
	Adj R^2	0.99993		0.99992			
Total students	Intercept	2867.04	0.0101	2980.941	0.0077	3044.891	0.0064
	W_STDT_TOT	0.007795	0.0255	0.008244	0.0187	0.007748	0.026
	POP_TOT	0.162	<.0001	0.161355	<.0001	0.161384	<.0001
	Adj R^2	0.99199		0.99187			

Table 3 (Continued)

Model	Variables	OLS Estimates	p-value	GS2SLS Estimates	p-value	GS3SLS Estimates	p-value
Out-commuters	Intercept	2500.981	0.5075	4573.762	0.3178	5416.765	0.0738
	W_COM_OUT	0.074225	0.0022	0.076689	0.0027	0.049511	0.0089
	POP_EAP	0.275318	<.0001	0.325594	<.0001	0.293916	<.0001
	EMP_TOT	-0.19169	<.0001	-0.25209	0.002	-0.18701	0.0001
	AREA	13.08926	0.0061	14.75031	0.0048	11.60425	0.0013
	A_EMP	-0.00015	<.0001	-0.00015	0.0001	-0.00019	<.0001
	C_EMP[a]	0.004082	0.0099	0.002917	0.1313	0.003467	0.0054
	EXP_ED	-0.13358	<.0001	-0.17435	0.0032	-0.14554	<.0001
	Adj R^2	0.91073		0.90294			
In-commuters	Intercept	-17248	<.0001	-16542.9	<.0001	-14894.8	<.0001
	W_COM_IN	0.022585	0.0824	0.028774	0.0403	0.041301	<.0001
	POP_EAP	-0.20717	<.0001	-0.30485	<.0001	-0.32454	<.0001
	EMP_TOT	0.714042	<.0001	0.847401	<.0001	0.90801	<.0001
	C_EMP	0.006158	0.0002	0.007596	<.0001	0.004096	<.0001
	AREA	8.128212	0.122	2.932154	0.6087	1.666451	0.7437
	A_EMP	-0.00023	<.0001	-0.00016	0.0001	-0.0002	<.0001
	Adj R^2	0.95433		0.94868			
Number of Firms	Intercept	7575.904	<.0001	5505.969	0.0002	5690.407	<.0001
	W_FIRM_TOT	-0.01497	<.0001	-0.0187	<.0001	-0.01549	<.0001
	POP_TOT	0.071471	<.0001	0.083347	<.0001	0.081191	<.0001
	AREA	-1.73317	0.378	1.971042	0.3525	1.2376	0.3895
	A_EMP	0.000083	<.0001	0.00004	0.0085	0.000048	<.0001
	Adj R^2	0.99347		0.99301			
Employment in non-basic sectors	Intercept	-38512.3	<.0001	-38509.1	<.0001	-39312.2	<.0001
	W_EMP_NBAS	0.002703	0.5473	0.001792	0.6901	-0.00042	0.9067
	EMP_TOT	0.86757	<.0001	0.867186	<.0001	0.815499	<.0001
	AREA	21.2381	0.008	21.21444	0.0081	16.74665	0.0051
	EXP_ED	0.231969	0.0009	0.232615	0.0008	0.256987	<.0001
	A_EXPED	-0.00034	<.0001	-0.00034	<.0001	-0.00027	<.0001
	C_EXPED	-0.00378	0.212	-0.0033	0.276	-0.00102	0.6404
	Adj R^2	0.9951		0.99509			

[a]Indicates that the significance of the coefficient change as we change estimation procedures.

Table 4. Generalized Spatial Three-Stage Least Squares Results using Different Weight Matrices

Model	Variables	Uniform		Weight and distance		Weight and distance squared	
		estimates	*p*-value	estimates	*p*-value	estimates	*p*-value
Local Revenue	Intercept	208901.9	<.0001	215844.7	<.0001	215844.7	<.0001
	W_REV_LOC	0.007976	0.6931	0.000245	0.9495	0.000245	0.9495
	POP_TOT	0.956666	<.0001	0.9272	<.0001	0.9272	<.0001
	EMP_NBAS	2.695301	<.0001	2.797318	<.0001	2.797318	<.0001
	COM_IN	-2.383	<.0001	-2.46465	<.0001	-2.46465	<.0001
Local Expenditure	Intercept	130199.4	<.0001	125231	<.0001	125231	<.0001
	W_EXP_LOC[a]	-0.01917	0.2649	-0.00759	0.0199	-0.00759	0.0199
	POP_TOT	0.755258	<.0001	0.828126	<.0001	0.828126	<.0001
	COM_IN	-2.9486	<.0001	-2.75825	<.0001	-2.75825	<.0001
	EMP_NBAS	2.937372	<.0001	2.696653	<.0001	2.696653	<.0001
Total Housing Units	Intercept	5016.323	<.0001	6234.474	<.0001	6234.474	<.0001
	W_HOUS_TOT[a]	0.022468	0.0031	0.002704	0.1291	0.002704	0.1291
	POP_TOT	0.279945	<.0001	0.278249	<.0001	0.278249	<.0001
	COM_IN	-0.4364	<.0001	-0.41414	<.0001	-0.41414	<.0001
Populatio n	Intercept	12724.93	<.0001	13857.61	<.0001	13857.61	<.0001
	W_POP_TOT	0.007379	0.0099	0.00118	0.0749	0.00118	0.0749
	POP_EAP	1.328597	<.0001	1.328956	<.0001	1.328956	<.0001
Economicall y Active Population	Intercept	-7554.63	<.0001	-8084.37	<.0001	-8084.37	<.0001
	W_POP_EAP[a]	-0.00288	0.0498	-0.00027	0.4188	-0.00027	0.4188
	POP_TOT	0.707171	<.0001	0.70867	<.0001	0.70867	<.0001
	EMP_NBAS	0.123211	<.0001	0.118865	<.0001	0.118865	<.0001
Students Total	Intercept	2391.439	0.0488	3044.891	0.0064	3044.891	0.0064
	W_STDT_TOT	0.03472	0.0136	0.007748	0.026	0.007748	0.026
	POP_TOT	0.161209	<.0001	0.161384	<.0001	0.161384	<.0001
Out-commuters	Intercept	-1535.2	0.5927	5416.765	0.0738	5416.765	0.0738
	W_COM_OUT	0.237803	<.0001	0.049511	0.0089	0.049511	0.0089
	POP_EAP	0.189387	<.0001	0.293916	<.0001	0.293916	<.0001
	EMP_TOT	-0.06215	0.1223	-0.18701	0.0001	-0.18701	0.0001
	AREA	10.68272	0.0025	11.60425	0.0013	11.60425	0.0013
	A_EMP	-0.00019	<.0001	-0.00019	<.0001	-0.00019	<.0001
	C_EMP	0.003746	0.0003	0.003467	0.0054	0.003467	0.0054
	EXP_ED	-0.06784	0.0066	-0.14554	<.0001	-0.14554	<.0001
In-commuters	Intercept	-16701.9	<.0001	-14894.8	<.0001	-14894.8	<.0001
	W_COM_IN	0.183798	<.0001	0.041301	<.0001	0.041301	<.0001
	POP_EAP	-0.33579	<.0001	-0.32454	<.0001	-0.32454	<.0001
	EMP_TOT	0.925822	<.0001	0.90801	<.0001	0.90801	<.0001
	C_EMP	0.003844	0.0001	0.004096	<.0001	0.004096	<.0001
	AREA	2.334738	0.6465	1.666451	0.7437	1.666451	0.7437
	A_EMP	-0.0002	<.0001	-0.0002	<.0001	-0.0002	<.0001

Table 4. (Continued)

Model	Variables	Uniform		Weight and distance		Weight and distance squared	
		estimates	*p*-value	estimates	*p*-value	estimates	*p*-value
Firms Total	Intercept	6436.575	<.0001	5690.407	<.0001	5690.407	<.0001
	W_FIRM_TOT	-0.06122	<.0001	-0.01549	<.0001	-0.01549	<.0001
	POP_TOT	0.079388	<.0001	0.081191	<.0001	0.081191	<.0001
	AREA	1.154086	0.4264	1.2376	0.3895	1.2376	0.3895
	A_EMP	0.000055	<.0001	0.000048	<.0001	0.000048	<.0001
Employment in non-basic sectors	Intercept	-35139.4	<.0001	-39312.2	<.0001	-39312.2	<.0001
	W_EMP_NBAS[a]	-0.06311	<.0001	-0.00042	0.9067	-0.00042	0.9067
	EMP_TOT	0.816538	<.0001	0.815499	<.0001	0.815499	<.0001
	AREA	14.66602	0.0137	16.74665	0.0051	16.74665	0.0051
	EXP_ED	0.266305	<.0001	0.256987	<.0001	0.256987	<.0001
	A_EXPED	-0.00029	<.0001	-0.00027	<.0001	-0.00027	<.0001
	C_EXPED[a]	0.00327	0.0756	-0.00102	0.6404	-0.00102	0.6404

[a]Indicates that the significance of the coefficient change as we change weight matrices.

Table 6. Economic Impact Estimated From Spatial Lag Model[a].

Province	Gun or Si	REV_LOC	EXP_LOC	HOUS_TOT	POP_TOT	POP_EAP	STDT_TOT	COM_OUT	COM_IN	FIRM_TOT	EMP_NBAS
Pusan	Gijang-Gun	3602.8	3202.5	481.87	2206.21	1660.13	356.00	266.77	319.16	191.57	813.28
Yulsan	Yulju-Gun	-7	-8.14	-1.52	-0.14	-0.09	-0.037	2.268	3.584	0.006	0.016
Gyung-Buk	Pohang-Si	-0.01	-0.01	-0.002	0	0	0	0.004	0.004	0	0
Gyung-Buk	Gyungju-Si	0.18	0.17	0.044	0.01	0	0.002	-0.07	-0.096	0.001	0
Gyung-Nam	Changwon-Si	0	0	-0.001	0	0	0	0	0.001	0	0
Gyung-Nam	Gimhae-Si	0.23	0.23	0.057	0.03	0.02	0.01	-0.008	-0.105	0.004	0.001
Gyung-Nam	Milyang-Si	0	0	-0.001	0	0	0	0	0.001	0	0
Gyung-Nam	Yangsan-Si	-6.98	-7.85	-1.57	-0.35	-0.24	-0.093	2.175	3.539	0.017	0.043
Pusan	Pusan	-6.69	-5	-2.096	-2.83	-1.93	-0.762	0.189	2.878	0.088	0.306
Yulsan	Yulsan	-6.95	-7.59	-1.623	-0.62	-0.42	-0.168	1.879	3.454	0.02	0.067
Gyung-Nam	Jinhae-Si	0	0	0	0	0	0	0	0.001	0	0
Total Impact		3575.58	3174.31	475.156	2202.31	1657.47	354.951	273.206	332.422	191.706	813.709

[a]Effects of 1,000 new jobs created in Gijang County of Punsan Province, Korea

Table 6b. Economic Impact Comparison of a Spatial and Non-Spatial Model[a]

Variable	Impact from spatial model	Impact from non-spatial model	Percentage difference
Local revenue (million won)	3603	3144	-13%
Local expenditures (million won)	3203	2587	-19%
Housing units	482	466	-3%
Population	2206	2253	2%
Economically active population	1660	1695	2%
Number of students	356	364	2%
Out-commuters	267	331	24%
In-commuters	319	381	19%
Number of firms	192	172	-10%
Employment in non-basic sector	813	830	2%

[a]Effects of 1,000 new jobs created in Gijang County of Punsan Province, Korea.

Table 7. Estimated Coefficients and Probability Value of Non-spatial, Spatial Error, Spatial Lag, and Spatial Lag and Error Models

Model	Variables	Non-spatial 3SLS		Spatial error 3SLS		Spatial lag 3SLS		Spatial lag and spatial error 3SLS	
		estimates	*p*-value	estimates	*p*-value	estimates	*p*-value	estimates	*p*-value
Local Revenues	Intercept	248958.7	<.0001	200378.1	<.0001	215844.7	<.0001	202998.7	<.0001
	W_REV_LOC					0.000245	0.9495	-0.00185	0.6453
	POP_TOT	0.514642	<.0001	0.906491	<.0001	0.9272	<.0001	0.920328	<.0001
	EMP_NBAS	4.331033	<.0001	2.904124	<.0001	2.797318	<.0001	2.733832	<.0001
	COM_IN	-4.22237	<.0001	-2.85821	<.0001	-2.46465	<.0001	-2.13439	<.0001
Local Expenditures	Intercept	161408.8	<.0001	144099.9	<.0001	125231	<.0001	130708.3	<.0001
	W_EXP_LOC					-0.00759	0.0199	-0.00744	0.0226
	POP_TOT	0.238276	0.0022	0.511064	<.0001	0.828126	<.0001	0.791542	<.0001
	COM_IN	-5.2472	<.0001	-3.70837	<.0001	-2.75825	<.0001	-2.6689	<.0001
	EMP_NBAS	4.881023	<.0001	3.745249	<.0001	2.696653	<.0001	2.77556	<.0001
Housing Units	Intercept	6752.514	<.0001	6945.258	<.0001	6234.474	<.0001	6283.123	<.0001
	W_HOUS_TOT					0.002704	0.1291	0.001779	0.323
	POP_TOT	0.279791	<.0001	0.270104	<.0001	0.278249	<.0001	0.277602	<.0001
	COM_IN	-0.43113	<.0001	-0.33298	<.0001	-0.41414	<.0001	-0.40934	<.0001
Population	Intercept	14879.16	<.0001	14897.17	<.0001	13857.61	<.0001	12421.59	<.0001
	W_POP_TOT					0.00118	0.0749	0.001284	0.0552
	POP_EAP	1.328957	<.0001	1.32289	<.0001	1.328956	<.0001	1.328152	<.0001
Economically Active Population	Intercept	-8539.31	<.0001	210.484	0.775	-8084.37	<.0001	-7945.61	<.0001
	W_POP_EAP[a]					-0.00027	0.4188	-0.00069	0.0547
	POP_TOT	0.712809	<.0001	0.58699	<.0001	0.70867	<.0001	0.722897	<.0001
	EMP_NBAS	0.107701	<.0001	0.445063	<.0001	0.118865	<.0001	0.080689	<.0001
Students Total	Intercept	4151.407	<.0001	5551.126	<.0001	3044.891	0.0064	3522.026	0.0015
	W_STDT_TOT					0.007748	0.026	0.006017	0.0757
	POP_TOT	0.161446	<.0001	0.15803	<.0001	0.161384	<.0001	0.16096	<.0001
Out-commuters	Intercept	12324.62	0.1878	3614.867	0.0994	5416.765	0.0738	2125.595	0.3771
	W_COM_OUT					0.049511	0.0089	0.049771	0.0128
	POP_EAP	0.479846	0.0213	0.24391	<.0001	0.293916	<.0001	0.225182	<.0001
	EMP_TOT	-0.4823	0.0582	-0.12963	<.0001	-0.18701	0.0001	-0.09711	0.0213
	AREA[a]	12.63693	0.1631	9.860337	0.0066	11.60425	0.0013	11.87638	0.0017
	A_EMP	-0.00017	0.0001	-0.00021	<.0001	-0.00019	<.0001	-0.00019	<.0001
	C_EMP[a]	0.004114	0.282	0.006234	<.0001	0.003467	0.0054	0.003854	0.0016
	EXP_ED[a]	-0.21599	0.2112	-0.09102	<.0001	-0.14554	<.0001	-0.10079	0.0006
In-commuters	Intercept	-12044.8	0.0007	-10000.9	0.0006	-14894.8	<.0001	-11242	0.0004
	W_COM_IN					0.041301	<.0001	0.03773	0.0005
	POP_EAP	-0.33594	<.0001	-0.28069	<.0001	-0.32454	<.0001	-0.35588	<.0001
	EMP_TOT	0.950525	<.0001	0.849605	<.0001	0.90801	<.0001	0.952104	<.0001
	C_EMP	0.00607	<.0001	0.006524	<.0001	0.004096	<.0001	0.004159	0.0002
	AREA	-1.04957	0.8397	2.165106	0.6726	1.666451	0.7437	-0.15282	0.9768
	A_EMP	-0.00024	<.0001	-0.00024	<.0001	-0.0002	<.0001	-0.00019	<.0001

Table 7 (Continued)

Model	Variables	Non-spatial 3SLS		Spatial error 3SLS		Spatial lag 3SLS		Spatial lag and spatial error 3SLS	
		estimates	*p*-value	estimates	*p*-value	estimates	*p*-value	estimates	*p*-value
Firms Total	Intercept	4119.179	<.0001	3321.61	0.0003	5690.407	<.0001	4311.7	<.0001
	W_FIRM_TOT					-0.01549	<.0001	-0.01479	<.0001
	POP_TOT	0.076313	<.0001	0.076157	<.0001	0.081191	<.0001	0.082411	<.0001
	AREA	1.765313	0.1887	1.719598	0.2136	1.2376	0.3895	1.778891	0.2243
	A_EMP	0.000066	<.0001	0.000072	<.0001	0.000048	<.0001	0.000045	<.0001
Employment in Non-basic Sectors	Intercept	-37640.7	<.0001	-24941.7	<.0001	-39312.2	<.0001	-17855.5	<.0001
	W_EMP_NBAS[a]					-0.00042	0.9067	-0.00667	0.0932
	EMP_TOT	0.829633	<.0001	0.866989	<.0001	0.815499	<.0001	0.842039	<.0001
	AREA[a]	14.56297	0.0115	13.19979	0.0111	16.74665	0.0051	4.091756	0.4368
	EXP_ED	0.22007	<.0001	0.160258	<.0001	0.256987	<.0001	0.174738	<.0001
	A_EXPED	-0.00024	<.0001	-0.00019	<.0001	-0.00027	<.0001	-0.00019	<.0001
	C_EXPED[a]	-0.00033	0.8019	0.008365	<.0001	-0.00102	0.6404	0.00058	0.7842

[a]Indicates that the significance of the coefficient change as we change estimation procedures.

Table 8. Mean Absolute Percentage Error as a Measure of Forecasting Accuracy in Different Models

Equations	Spatial lag and spatial error model	Spatial error model	Spatial lag model	Non-spatial model
REV_LOC	21.8	22.5	21.9	25.1
EXP_LOC	15.0	15.6	15.2	19.1
HOUS_TOT	13.0	13.7	12.7	13.6
POP_TOT	11.0	11.9	11.1	11.9
POP_EAP	10.2	9.1	9.1	9.7
STDT_TOT	48.4	53.0	45.5	53.3
COM_OUT	423.0	544.7	470.2	730.4
COM_IN	173.0	167.4	177.6	183.5
FIRM_TOT	21.3	22.4	23.4	22.6
EMP_NBAS	72.2	86.0	85.8	87.1
Average	81.0	94.7	87.3	115.8
Coefficient of variation	4.0	4.6	4.2	5.1

REFERENCES

Anselin, L. 1988. *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer Academic Publishers.

Berck, P. and K. Sydsaeter. 1993. *Economist's Mathematical Manual, 2nd Edition*. Berlin: Springer-Verlag.

Case, A.C., J.R. Hines, and H.S. Rosen. 1993. Budget spillovers and fiscal policy interdependence: Evidence from the states. *Journal of Public Economics* 52: 285-307.

Cliff, A.D. and J.K. Ord. 1973. *Spatial Autocorrelation.* London: Pion.

Cooke, S., and L. Fox. "Using the Idaho Fiscal Impact Model for Local Fiscal Impact Assessment." *Agricultural Economics Research Series 94-18,* Dept. Agricultural Economics and Rural Sociology, University of Idaho, Moscow, Idaho, October, 1994.

Evans, G.K. and J.I. Stallmann. 2006. SAFESIM: The Small Area Fiscal Estimation Simulator. In T.G. Johnson, D.M. Otto and S.C. Deller (eds.), *Community Policy Analysis Modeling.* Ames, IA: Blackwell Publishing.

Johnson, T.G. 1991. A description of the VIP model. Unpublished manuscript, Department of Agricultural Economics, Virginia Polytechnic Institute and State University, Blacksburg.

Johnson, T.G. and J.K. Scott. 2006. The Show Me Community Policy Analysis Model. In T.G. Johnson, D.M. Otto and S.C. Deller (eds.), *Community Policy Analysis Modeling* (pp.119-129). Ames, IA: Blackwell Publishing.

Kelejian, H.H. and I.R. Prucha. 2004. Estimation of simultaneous systems of spatially interrelated cross sectional equations. *Journal of Econometrics* 118: 27-50.

Kelejian, H. H. and D. Robinson. 1995. Spatial correlation: A suggested alternative to the autoregressive model. In *New Directions in Spatial Econometrics.* New York: Springer-Verlag.

Murdoch J.C., M. Rahmatian, and M.A. Thayer. 1993. A spatially autoregressive median voter model of recreation expenditures. *Public Finance Quarterly* 21: 334-350.

Shields, M. 1998. An integrated economic impact and simulation model for Wisconsin counties. Unpublished Doctoral Dissertation, University of Wisconsin—Madison.

Swenson, D. and D. Otto. 2000. The Iowa Economic/Fiscal Impact Modeling System. *Journal of Regional Analysis and Policy,* 28:64-75.

Tiebout, C.M. 1956. A pure theory of local expenditures. *The Journal of Political Economy* 64(5): 416-424.