Estimation of Simultaneous System of Spatial Dynamic Equation Models with Spatially Corrected Blundell Bond Method of Moment (SCBB-GMM)

Faiqotul Mala*, Rahma Fitriani¹, Nurjannah¹
¹Statistic Department, Faculty of Natural Sciences, University of Brawijaya

*Corresponding author: faiqmalaa@gmail.com¹

Abstract. The purpose of this paper is used to estimate spatial dynamic simultaneous equations models with simultaneous effects, spatial effects, and time lag effects. Spatial interactions are presented by spatial weight matrices and in addition to dynamics in space and time. Many variables commonly have two direction relationships. Two directions relationship of mutual influence can be summarized in a system of simultaneous equations. In other ways, the relationships of variables are dynamic. In the dynamic panel data models, there is a lag of dependent variable that caused the OLS estimates to be biased and inconsistent. First difference and instrumental variable are needed to overcome this problem. Estimation of parameters in spatial dynamic panel data models with Spatially Corrected Blundell Bond using principles of Generalized Method of Moment (GMM). SCBB-GMM with optimal weights produce parameters that are unbiased, consistent, and efficient

Keywords: Spatially Corrected Blundell Bond, GMM, Spatial Dynamic Panel Data, Simulant Equation, Instrumental Variable.

1. Introduction
The single model more common in used is to ignore the interdependence between variables. In the case of the economy, many variables commonly have a two-way relationship. Two-way relationship of mutual influence can be summarized in a system of simultaneous equations. Almost all approaches in macroeconomics have simultaneous properties [1]. The simultaneous system consists of some equations, where the sum of the equation is equal to the sum of the value variable to explain that is called endogenous variable. The contribution of some variable to the model is called by a predetermined variable. However, this variable consists of the exogenous variable and the lagged endogenous variable.

In the expand, simultaneous equation models included an aspect of interrelationships between locations in the model. A variable in the model is not only caused by explanatory variables but also caused by spatial effects or spatial interactions between locations. There are three kinds of effects of spatial interactions (1) included spatial interactions between dependent variables, (2) spatial interactions between explanatory variables, and (3) spatial interactions between error terms [2].
In the economic variables, it is not enough to use cross-section data because it is needed to observe the individual effects at various time periods. Data which is a combination of cross-section data and time series data is called panel data. There are several benefits to using panel data, data is heterogeneous, more informative, varied, the degree of freedom is large, more efficient, dynamic, enable to detect factors are not observed in cross-section data and time series, and minimize bias. In addition, the relationship of economic variables is dynamic, which is not only influenced by variables at the same time but influenced by variables in the past [3]. To estimate the parameters in dynamic panel spatial structural equations using SCBB-GMM which produces an unbiased, consistent, and efficient estimator.

2. Model

2.1. Simultaneous Equations Models

The general simultaneous equations with \( M \) endogenous variables denoted \( (Y_1, Y_2, \ldots, Y_M) \) and \( K \) exogenous variables \( (X_1, X_2, \ldots, X_K) \) may be written as:

\[
\begin{align*}
\alpha_{11} Y_{1i} + \alpha_{12} Y_{2i} + \cdots + \alpha_{1M} Y_{Mi} + \beta_{11} X_{1i} + \beta_{12} X_{2i} + \cdots + \beta_{1k} X_{ki} &= \epsilon_{1i}, \\
\alpha_{21} Y_{1i} + \alpha_{22} Y_{2i} + \cdots + \alpha_{2M} Y_{Mi} + \beta_{21} X_{1i} + \beta_{22} X_{2i} + \cdots + \beta_{2k} X_{ki} &= \epsilon_{2i}, \\
\vdots & \vdots \\
\alpha_{Mi} Y_{1i} + \alpha_{M2} Y_{2i} + \cdots + \alpha_{MM} Y_{Mi} + \beta_{M1} X_{1i} + \beta_{M2} X_{2i} + \cdots + \beta_{Mk} X_{ki} &= \epsilon_{mi},
\end{align*}
\]

(1)

with \( \epsilon_{1i}, \epsilon_{2i}, \ldots, \epsilon_{mi} \) is structural disturbance \( \alpha_{ij} \) is coefficients of the \( j \)th endogenous variables and the \( j \)th equations with \( j = 1,2, \ldots, M \) and \( \beta_{jl} \) is coefficients of the predetermined variables with \( j = 1,2, \ldots, M \) and \( l = 1,2, \ldots, K \), while \( i \) total number of observations, \( i = 1,2, \ldots, n \).

Equation (1) can be written as matrix:

\[
\begin{bmatrix}
\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1M} & Y_{1i} & \beta_{11} & \beta_{12} & \cdots & \beta_{1k} & X_{1i} & \epsilon_{1i} \\
\alpha_{21} & \alpha_{22} & \cdots & \alpha_{2M} & Y_{2i} & \beta_{21} & \beta_{22} & \cdots & \beta_{2k} & X_{2i} & \epsilon_{2i} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
\alpha_{Mi} & \alpha_{M2} & \cdots & \alpha_{MM} & Y_{Mi} & \beta_{M1} & \beta_{M2} & \cdots & \beta_{Mk} & X_{Mi} & \epsilon_{Mi}
\end{bmatrix}
\]

(2)

\[
\begin{bmatrix}
Y_{1i} \\
Y_{2i} \\
\vdots \\
Y_{Mi}
\end{bmatrix}
= \begin{bmatrix}
\beta_{11} & \beta_{12} & \cdots & \beta_{1k} \\
\beta_{21} & \beta_{22} & \cdots & \beta_{2k} \\
\vdots & \vdots & \ddots & \vdots \\
\beta_{M1} & \beta_{M2} & \cdots & \beta_{Mk}
\end{bmatrix}
\begin{bmatrix}
X_{1i} \\
X_{2i} \\
\vdots \\
X_{Mi}
\end{bmatrix}
+ \begin{bmatrix}
\epsilon_{1i} \\
\epsilon_{2i} \\
\vdots \\
\epsilon_{Mi}
\end{bmatrix}
\]

(3)

The solution of the system of equations determining \( Y_j \) in terms of \( X_j \) and \( \epsilon_j \) is the reduced form of the model,

\[
Y_j = \Pi X_j + v_j
\]

By the identification problem, it means whether numerical estimates of the parameters of a structural equation can be obtained from the estimated reduced-form coefficients [1]. An identified equation may be either exactly identified or overidentified. It is said to be exactly identified if unique numerical values of the structural parameters can be obtained. It is said to be overidentified if more than one numerical value can be obtained for some of the parameters of the structural equations. To identified simultaneous equations there is have two conditions, the first is order condition and the second is the rank condition. The notation to identification can be written by

- \( M \): the sum of endogenous variables in the model
- \( m \): the sum of endogenous variables in the structural equation
- \( K \): some of the predetermined variables in the model
- \( k \): some of the predetermined variables in the structural equation.

Identification of the simultaneous equation has two conditions, for the first is exactly identified when \( K - k = m - 1 \); \( \text{rank} \ A = M - 1 \) and overidentified when \( K - k > m - 1 \); \( \text{rank} \ A = M - 1 \). The Ordinary Least Squares (OLS) estimators of the structural parameters are inconsistent and biased
because of the included endogenous variables in each equation are correlated with the disturbances. Alternatively, used Two-Stage Least Square (2SLS) method for estimate parameter. The 2SLS estimator is obtained by OLS regression of $Y_j$ on $\mathbf{Y}$ and $X_j$. The two-stage regressions in the procedure are to obtain the least squares predictions from regression of $Y_j$ on $X_j$, and estimate the parameter by least square regression of $Y_j$ on $\hat{Y}_j$ and $X_j$. To prove that the equations have simultaneity used Hausman test [4].

2.2. Spatial Dynamic Panel Model

Consider a panel with $i=1,2,...,N$ spatial units and $t=1,2,...,T$ time periods. Assume that the data at time $t$ are generated according to the following model:

$$y_t = \rho W y_t + \gamma y_{t-1} + \delta y_{t-1} + X_t \beta + u_t$$  \hspace{1cm} (4)

where $y_t$ is an $N \times 1$ vector of observations on the dependent variable, $y_{t-1}$ is a one period time lag of the dependent variable, $W$ is an $N \times N$ matrix of spatial weight, $X_t$ is an $N \times K$ matrices of observations on the strictly exogenous explanatory variables, $u_t$ is an $N \times 1$ vector of error terms, $\rho$ is the spatial autoregressive coefficient, $\delta$ is the coefficient of lagged dependent variable, $\gamma$ is the endogenous explanatory coefficient and $\beta$ is a $K \times 1$ vector exogenous coefficients [5].

3. Spatial Dynamic Panel Estimators

In this section, the spatial dynamic panel estimators are proposed with extended the static panel data model of Kapoor et al. (2007) [6]. Because there is a time lag of the dependent variable, for estimation apply GMM procedure. Propose sets of instruments for both the time lag and spatial lag of the dependent variable. This procedure yields consistent Spatially Corrected Blundell-Bond estimators, which will be derived in two stages [5].

3.1. Blundell-Bond Estimator

In the dynamic model, there is endogeneity which is dependent variable correlated with the error $E(x_{it}, \varepsilon_{it}) \neq 0$ or $cov(x_{it}, \varepsilon_{it}) \neq 0$ and OLS estimator will be biased and inconsistent. Anderson and Hsiao (1982) introduce method Instrumental Variable (IV) with build other variables which is not correlated with error but correlated with endogenous explanatory. Arellano and Bond (1991) yield the first difference unbiased and consistent estimator called AB-GMM.

The standard Arellano-Bond estimator is known to be rather inefficient when instruments are weak because it makes use of the information contained in the first differences of variables only. To address this shortcoming, the GMM approach of Blundell and Bond's (1998) referred to the system GMM estimator. Extends the Arellano and Bond (1991) conditions by specification moment conditions also for variables in levels which consist of both first differenced and level equation and an extended set of instruments variable. For the case included the spatial effect, the parameter is estimated by Spatially Corrected Blundell-Bond Generalized Method of Moment (SCBB-GMM) [5].

3.1.1 The First Stage. The Blundell-Bond estimator for the spatially autoregressive dynamic panel model can be derived by stacking equation (4) and:

$$Y_t = Q_t \theta + u_t$$  \hspace{1cm} (5)

with $Q_t = [Y_0, W Y_0, Y_{t-1}, X_t]$, $\theta = [\rho_j, \gamma_j, \delta_j, \beta_j]$, $y_t \neq Y_t$

Then, take first differences of (5):

$$(y_t - y_{t-1}) = Q \theta + (u_t - u_{t-1})$$  \hspace{1cm} (6)
can be written

\[ \Delta y_t = Q_t \theta + \Delta u_t \]  

(7)

To estimate \( \theta \), it followed a GMM estimator by a set of linear moment conditions for the error term. Consistent GMM estimation is possible if there are some of instruments are correlated with the time lagged, spatially lagged, and exogenous variables and are uncorrelated with the errors \( \nu_t \) for each \( t = 3, \ldots, T \). The moment condition first difference models are

\[
E(W \Delta Y_{t}, \nu_t) = 0 ; t = 3, \ldots, T \\
E(\Delta Y_t, \nu_t) = 0 ; t = 3, \ldots, T \\
E(\Delta Y_{t-s}, \nu_t) = 0 ; t = 3, \ldots, T; s = 2, \ldots, t - 1 \\
E(\Delta X_t, \nu_t) = 0 ; t = 3, \ldots, T
\]

The matrix of instruments is defined as \( (Z_{\text{diff}}) \) has the following structure

\[
Z_{\text{diff}} = \begin{bmatrix}
y_{ij1} W y_{ij1} & 0 & 0 \\
0 & y_{ij1} y_{ij2} W y_{ij1} W y_{ij2} & 0 \\
0 & 0 & y_{ij1} y_{ij2} \ldots y_{ij(t-2)} \ldots W y_{ij1} \ldots W y_{ij(t-2)}
\end{bmatrix}
\]

Then construct set the moment condition for the model level

\[
E(W Y_t', \Delta \nu_t) = 0 ; t = 3, \ldots, T \\
E(Y'_t, \Delta \nu_t) = 0 ; t = 3, \ldots, T \\
E(Y'_{t-s}, \Delta \nu_t) = 0 ; t = 3, \ldots, T; s = 2, \ldots, t - 1 \\
E(\Delta X_t, \Delta \nu_t) = 0 ; t = 3, \ldots, T
\]

The matrix of instruments is defined as \( (Z_{\text{lev}}) \) has the following structure

\[
Z_{\text{lev}} = \begin{bmatrix}
\Delta y_{ij2} W \Delta y_{ij2} & 0 & 0 \\
0 & \Delta y_{ij2}, \Delta y_{ij3} W \Delta y_{ij2}, \Delta y_{ij3} & 0 \\
0 & 0 & \Delta y_{ij2}, \Delta y_{ij2}, \ldots W \Delta y_{ij2}, \ldots W \Delta y_{ij2}
\end{bmatrix}
\]

\( Z_{\text{diff}} \) and \( Z_{\text{lev}} \) are combined into a matrix instrument variable for SCBB-GMM

\[
Z = \begin{bmatrix}
Z_{\text{diff}} \\
Z_{\text{lev}}
\end{bmatrix}
\]

(8)

Then construct the estimator model as follows

\[
\begin{bmatrix}
\Delta Y_t \\
Y_t
\end{bmatrix} = \begin{bmatrix}
\Delta Q_t \\
Q_t
\end{bmatrix} \theta + \begin{bmatrix}
\Delta \nu_t \\
\nu_t
\end{bmatrix}
\]

(9)
Equation (9) can be written by

\[ y_i = \theta Q + u_t \]  

(10)

In the first stage, the SCBB-GMM estimator is derived:

\[ \hat{\theta} = [Q' \hat{Z} \hat{W} Q]^{-1} [Q' \hat{Z} \hat{W} Y] \]  

(11)

3.1.2 The Second Stage. Formulate the spatial weighted matrix \( W \), which is an unbiased and consistent estimator for \( W_{(L \times L)} \), where \( L \) is the sum of instrument variable. Blundell and Bond (1998) propose optimal weight matrix

\[ \Psi = N^{-1} \sum_{i=1}^{N} Z^i u \]  

(12)

Substitution equation (12) to (11), the estimator of \( \theta \) in the final stage is

\[ \hat{\theta} = [Q' \Psi Z' Q]^{-1} [Q' \Psi Z' y] \]  

(13)

4. Result and Discussion

The simultaneous system model is a model that endogenous variable can be exogenous variable in another side. Estimation method 2SLS needs a reduced model to estimate reduced form for the first. In the other hand, the reduced form is a data panel regression and consist of variable exogenous. The structural equation for simultaneous equation models can be noted by:

\[ y_t = \rho W Y_t + \gamma Y_t + \delta Y_{t-1} + X_t \beta + u_t \]

with \( Q_t = [Y_t, W Y_t, Y_{t-1}, X_t] \), \( \theta = [\rho^', \gamma^', \delta^', \beta^'] \), \( y_t \neq Y_t \)

Basis problem in the model data panel is lagged endogenous variable correlated with error term, OLS estimator is bias and inconsistent. So that SCBB-GMM estimator is used because the parameter will estimate is unbias, consistent and efficient. First, take first differences of (5):

\[ (y_t - y_{t-1}) = Q \theta + (u_t - u_{t-1}) \]

\[ \Delta y_t = Q_t \theta + \Delta u_t \]

\( Z_{diff} \) and \( Z_{lev} \) are combined into a matrix instrument variable for SCBB-GMM

\[ Z = \begin{bmatrix} Z_{diff} & 0 \\ 0 & Z_{lev} \end{bmatrix} \]

Then construct the estimator model as follows

\[ \begin{bmatrix} \Delta y_t \\ y_t \end{bmatrix} = \begin{bmatrix} \Delta Q_t \\ Q_t \end{bmatrix} \theta + \begin{bmatrix} \Delta v_t \\ v_t \end{bmatrix} \]
\[ y_t = Q\theta + u_t \]
\[ u_t = y_t - Q\theta \]

Formulated moments of population conditions
\[ E(g_i(\theta)) = E(Z_i'v_i) = E(Z_i'(y_t - Q\theta)) = 0 \]

Formulated sample moments conditions
\[ g(\theta) = N^{-1}\sum_{i=1}^{N} Z_i(y_t - Q\theta) \]

Build the GMM function from moments of sample conditions
\[ J(\theta) = g(\theta)'\Psi g(\theta) \]

\[ \hat{\theta} \] estimated with take first derivative \( J(\hat{\theta}) \)
\[ \frac{\partial J(\theta)}{\partial \theta} = 0 \]
\[ \hat{\theta} = \left[ \sum_{i=1}^{N} (Q'Z_i) \right]^{-1} \left[ \sum_{i=1}^{N} Z_i'Z_i - \hat{\theta}^2 \right] \]

Formulate the optimal weight matrix
\[ \Psi = N^{-1}\sum_{i=1}^{N} Z_i'Z_i \]

\[ \Psi = N^{-1}\sum_{i=1}^{N} Z_i'Z_i \Theta - \hat{\theta}^2 \]

the estimator of \( \theta \) in the final stage is
\[ \hat{\theta} = \left[ \sum_{i=1}^{N} (Q'Z_i) \right]^{-1} \left[ \sum_{i=1}^{N} Z_i'Z_i \right] \]

\[ \Theta = \left[ Q'Z\Psi Z'Q \right]^{-1} [Q'Z\Psi Z'y] \]

5. Conclusion
The estimation process in a simultaneous system model with dynamic panel data is using the 2SLS method, start from estimation in the reduced form then estimating the full model (all of the structural equations). Both of process estimated by SCBB-GMM in two-stages. The method is can be used in the economics case, where the relationship between variables are dynamic and there is a spatial effect.
References

[1] Greene, W. H. 2003. Econometrics Analysis. 5th edition. Prentice Hall. New Jersey.

[2] Manski, C.F. 1993. Identification of Endogenous Social Effects: The Reflection Problem. The Review of Economic Studies, 60(3), 531-542.

[3] Baltagi, B. H. 2005. Econometric Analysis of Panel Data. New York: John Wiley & Sons.

[4] Greene, W. H. 2003. Econometric Analysis. 5th edition. Prentice Hall. New Jersey.

[5] Cizek P, Jacobs J, Lighart J dan Vrijburg. 2011. GMM Estimation of fixed Effects Dynamic Panel Data Models with Spatial Lag and Spatial Errors. Journal of Econometrics 150 : 97-134.

[6] Kapoor M, Kelejian HH, Prucha IR. 2007. Panel Data Models With Spatially Correlated Error Components. Journal of Econometrics 140: 97-130.