Simultaneous equation models with spatially autocorrelated error components

AMBA OYON, Claude Marius and Mbratana, Taoufiki

University of Yaounde II, University of Yaounde II

October 2017

Simultaneous Equation Models with Spatially Autocorrelated Error Components

Marius Claude Oyon Amba[*]
Taoufiki Mbratana[†]

October 2017

Abstract
This paper develops estimators for simultaneous equations with spatial autoregressive or spatial moving average error components. We derive a limited information estimator and a full information estimator. We give the generalized method of moments to get each coefficient of the spatial dependence of each equation in spatial autoregressive case as well as spatial moving average case. The results of our Monte Carlo suggest that our estimators are consistent. When we estimate the coefficient of spatial dependence it seems better to use instrumental variables estimator that takes into account simultaneity. We also apply these set of estimators on real data.

Keywords: Panel data, SAR process, SMA process, Simultaneous equations, Spatial error components.
JEL: C13, C33.

[*]Corresponding author: Department of Quantitative techniques, University of Yaounde II, Cameroon. Tel:(237)656-082-448. Email: mariusamba@gmail.com

[†]Faculty of Economics and Managment, University of Yaounde II, Cameroon. Email: tao.mbratana@gmail.com

1 Introduction

Recently there has been increasing consideration in models with spatial interactions. Spatial models have received substantial interest in traditional econometrics as well, both from a theoretical as well as from an applied perspective, as suggested by the growing number of studies that are using spatial methods. In view of that, nowadays spatial regression methods are becoming an important part of the toolbox of applied econometrics and the interest is increasingly shifting away from the single-equation cross-sectional background to more sophisticated settings such as panel data models, qualitative variables models, simultaneous models or multilevel models in a spatial context (see Anselin, 2006, for a related literature review).

On the other hand, one of the most widely used spatial models is the spatial autoregressive (SAR) model based on a single equation introduced by Cliff and Ord (1973) and Cliff and Ord (1981)[1] . This Cliff-Ord type model has recently received substantial attention in various fields of economics (health, labour and public economics, political economy, international and urban economics) as it provides a suitable framework to model the interaction between economic agents[2]. Most of the literature focused on single-equation models where a single dependent variable is determined for cross sectional units. However, in economics it is recurrent that the outcomes for many dependent variables are determined jointly by a system of equations for units. In this situation, the simultaneous nature of the outcomes can arise from two sources, interactions between different economic variables as well as interactions between cross sectional units.

Somewhat surprisingly, the literature on the estimation of simultaneous systems of spatially interrelated cross sectional equations has so far been limited with some exceptions. Kelejian and Prucha (2004) extend the methodology developed in Kelejian and Prucha (1998) and Kelejian and Prucha (1999) for single equations, an early development of generalized method of moments (GMM) estimators for the simultaneous equation SAR model. They propose both limited information two stage least squares (2SLS) and full information three stage least squares (3SLS) estimators and derive for these estimators their asymptotic properties. Liu (2014) and Zenou (2017) exploit the methodology of Kelejian and Prucha (2004) within the context of social interaction models, and provide further refinements. Other recent contribution to the literature on spatial simultaneous equation models are Wang et al. (2014) who analyse the quasi maximum likelihood (QML) estimator for such a system in the cross section. Prucha et al. (2016) developed an estimation methodology for network data generated from a system of simultaneous equations. Their specification allows for network interdependencies via spatial lags in the endogenous and exogenous variables, as well as in the disturbances. By allowing for higher-order spatial lags, their specification provides important flexibility in modeling network interactions. For a simultaneous equation SAR model, Liu and Saraiva (2017) provided a GMM estimator and its heteroskedasticity-robust

[1]This model is a variant of the model introduced by Whittle (1954) and is sometimes referred to as a spatial autoregressive model; see, e.g., Anselin (1988).

[2]Early development in estimation and testing for cross sectional data can be found in Anselin (1988), Kelejian and Robinson (1992), Anselin and Bera (1998), and Cressie (2015), among others.

standard error. They established the consistency and asymptotic normality of the proposed GMM estimator and also show that it performs well in finite samples. Yang and Lee (2017) studied parameter spaces, parameter identification and asymptotic properties of the QML estimation in the framework of the simultaneous equation SAR model which includes simultaneity effects, own-variable spatial lags and cross-variable spatial lags as explanatory variables, and allows for correlation between disturbances across equations. Their main findings reveal that the QML estimator is asymptotically more efficient than the 3SLS estimator as the former implicitly uses additional information on the covariance structure of model disturbances. The authors also discussed a multivariate SAR model that can be considered as a reduced form of the simultaneous equations model.

Furthermore, the studies on spatial simultaneous equations model empirically have been motivated: see, Ho and Hite (2008); Jeanty et al. (2010); Allers and Elhorst (2011); Gebremariam et al. (2011); Baltagi and Bresson (2011); De Graaff et al. (2012); Hauptmeier et al. (2012); Goldsmith-Pinkham and Imbens (2013), among others.
Another strand of the literature focused on simultaneous spatial panel data models. Although the panel data simultaneous equations models that ignored the spatial autocorrelation have been developed (see, e.g., Baltagi, 1981; Prucha, 1985; Balestra and Varadharajan-Krishnakumar, 1987; Cornwell et al., 1992; Baltagi and Li, 1992), the simultaneous panel data models including spatial dependence structures are practically absent from the econometrics literature, with the possible exceptions of Baltagi and Deng (2015) and Lu (2017). In the context of the Kelejian and Prucha (1998) and Lee (2003) type instruments and the Baltagi (1981) error components 3SLS estimator, Baltagi and Deng (2015) propose a spatial error component 3SLS (SEC-3SLS) system estimator that handles endogeneity, spatial lag dependence, random effects as well as cross equation correlation simultaneously. Lu (2017) considered a simultaneous spatial panel data model, jointly modeling three effects, namely simultaneous effects, spatial effects and common shock effects and proposes the QML and the iterative generalized principal components (IGPC) methods to estimate the model. For each method, she determined its identification condition and developed a full inferential theory for its estimators and found that the estimators from both methods are consistent.

This paper extends the Baltagi (1981) EC-3SLS panel data estimator that ignores spatial dependence. This more general model allows for correlation across space, time, and equations. It combines the simplicity of dealing with heterogeneity in the panel using an error component model and spatial correlation disturbances. Besides, a well-known feature of the SAR specification is that it allows for a global transmission of shocks through global spillovers that agglomerate from higher order neighbours (Anselin, 1988; LeSage and Pace, 2009). Contrariwise, the SAR process may not be suitable, if the shocks are not transmitted globally. Therefore, another specification that allows for a localized transmission of shocks is required. Haining (1978), Anselin (1988), and recently Hepple (2003), Fingleton (2008), Baltagi and Pirotte (2011), Doğan and Taşpınar (2013), consider a spatial moving average (SMA) process for the disturbances. Consequently in the scope of this study, the spatial dependence specifications for the disturbance term uses a SAR as well as a SMA disturbances. For this purpose, the rest of the paper is organized as follow. Section 2 completly defines the model; section 3 deals with the limited information estimator; section 4 resolve the problem of estimation of spatial component ρ of each equation both in the SAR case and the SMA case; section 5 deals with the full information estimator; In section 6 we regroup a battery of simulations in model's parameters; section 7 applies theses estimators on real data, while in section 8 we conclude the paper.

2 The model

Consider the following lth structural equation

$$y_l = Y_l\alpha_l + X_l\beta_l + \upsilon_l = Z_l\delta_l + \upsilon_l \quad l = 1, \dots, L \tag{1}$$

where, $y_{l=1,\dots,L}$ are endogenous variables with dimension $TN \times 1$, Y_l is the set of $TN \times (M_l - 1)$[3] right-hand side endogenous variables included in equation l, X_l is a $TN \times K_l$ of right-hand side exogenous variables of the model included in the equation and $Z_l = [Y_l, X_l]$ is a matrix of explanatory variables of the equation; $\delta_l' = (\alpha_l', \beta_l')$ are the associate coefficients of Z_l.
We assume that the disturbances are generated either by a spatially autoregressive (SAR) process or a spatially moving average (SMA) process:

$$\upsilon_l = \Lambda_l \epsilon_l = \begin{cases} \rho_l W_l \upsilon_l + \epsilon_l = (I - \rho_l W_l)^{-1}\epsilon_l & \text{for SAR} \\ \rho_l W_l \epsilon_l + \epsilon_l = (I + \rho_l W_l)\epsilon_l & \text{for SMA} \end{cases} \tag{2}$$

In this section, $i = 1, \dots, N$ and $t = 1, \dots, T$; where N denotes the number of individuals and T the number of time periods. We order the observations first by time and then individuals because this grouping is more convenient for modelling spatial correlation via equation (2). $W_l = I_T \otimes W_{lN}$ with I_T being an identity matrix of dimension T and W_{lN} being a $N \times N$ spatial weighting matrix of known constants which does not involve time and is usually row-normalized. For $1 \leq l \leq L$, all diagonal elements of W_{lN} are zero. $|\rho_l| < 1$ is a scalar autoregressive parameter, and ϵ_l is a $TN \times 1$ vector of innovations.
To allow for the innovations to be correlated over time, we assume the following error component structure for the innovation vector ϵ_l

$$\epsilon_l = Z_\eta \eta_l + \xi_l \tag{3}$$

where $Z_\eta = \iota_T \otimes I_N$, $\eta_l' = \begin{pmatrix} \eta_{1l} & \dots & \eta_{Nl} \end{pmatrix}'$ represents the vector of unit specific error components; and $\xi_l' = \begin{pmatrix} \xi_{11l} & \xi_{12l} & \dots & \eta_{TNl} \end{pmatrix}'$ contains the error components that vary over both the cross-sectional units and time periods; ι_T is a $T \times 1$ vector of ones. η_l and ξ_l are centered random vector with covariance matrix

$$\mathrm{E}\begin{pmatrix} \eta_l \\ \xi_l \end{pmatrix} \begin{pmatrix} \eta_q' & \xi_q' \end{pmatrix} = \begin{bmatrix} \sigma^2_{\eta_{lq}} I_N & 0 \\ 0 & \sigma^2_{\xi_{lq}} I_{TN} \end{bmatrix} \tag{4}$$

[3]If we let M_{-l} the set of right-hand side endogenous variables excluded of the lth, we have $M_l + M_{-l} = L$.

In light of equation (3), the covariance matrix $\Omega_{\epsilon_{lq}} = \mathrm{E}(\epsilon_l \epsilon_q')$ between the lth and the qth equation is:

$$\Omega_{\epsilon_{lq}} = \sigma^2_{\eta_{lq}}(J_T \otimes I_N) + \sigma^2_{\xi_{lq}} I_{TN} \quad for \quad l, q = 1, \ldots, L \tag{5}$$

and its spectral decomposition can be written in compact form as below:

$$\Omega_{\epsilon_{lq}} = \sum_{h=0}^{1} \sigma^2_{h_{lq}} Q_h \quad for \quad l, q = 1, \ldots, L \tag{6}$$

where,

$$Q_h = B_h \otimes I_N,$$

$B_{h,\{h=0,1\}}$ is a square matrix of order T with $B_0 = E_T$ and $B_1 = \bar{J}_T$; $\bar{J}_T = \iota_T \iota_T'/T$ denoting average matrix over time, $E_T = I_T - \bar{J}_T$, $\sigma^2_{0_{lq}} = \sigma^2_{\xi_{lq}}$ and $\sigma^2_{1_{lq}} = T\sigma^2_{\eta_{lq}} + \sigma^2_{\xi_{lq}}$. The matrices Q_0 and Q_1 are standard transformation matrices utilized in the error component literature, with the appropriate adjustments implied by our adopted ordering of the data; compare, e.g., Baltagi (2008). They are symmetric, idempotent and orthogonal to each other. Furthermore, by letting $\mathrm{tr}(M)$ the trace of a square matrix M,

$$Q_0 + Q_1 = I_{TN}, \ \mathrm{tr}(Q_0) = (T-1)N \text{ and } \mathrm{tr}(Q_1) = N$$

and the covariance matrix of $\Omega_{\upsilon_{lq}}$ between the lth and the qth equation is:

$$\begin{aligned} \Omega_{\upsilon_{lq}} &= \Lambda_l \mathrm{E}(\epsilon_l \epsilon_q') \Lambda_q' \\ &= \Lambda_l (\sigma^2_{0_{lq}} Q_0 + \sigma^2_{1_{lq}} Q_1) \Lambda_q' \\ &= \sigma^2_{0_{lq}} \Lambda_l Q_0 \Lambda_q' + \sigma^2_{1_{lq}} \Lambda_l Q_1 \Lambda_q' \\ &= \sigma^2_{0_{lq}} Q_0 \Lambda_l \Lambda_q' + \sigma^2_{1_{lq}} Q_1 \Lambda_l \Lambda_q' \\ &= \sigma^2_{0_{lq}} Q_0 \Lambda_{lq} + \sigma^2_{1_{lq}} Q_1 \Lambda_{lq} \end{aligned} \tag{7}$$

where $\Lambda_{lq} = \Lambda_l \Lambda_q'$ and Λ_l non-singular $\forall l$; This comes from the fact that, for SMA process

$$\begin{aligned} W_l Q_h &= (I_T \otimes W_{lN})(B_h \otimes I_N) \\ &= (I_T B_h) \otimes (W_{lN} I_N) \\ &= (B_h I_T) \otimes (I_N W_{lN}) \\ &= Q_h W_l \end{aligned}$$

while for SAR process, since each $|\rho_l| < 1$, we have

$$
\begin{aligned}
\Lambda_l Q_h &= (I - \rho_l W_l)^{-1} Q_h \\
&= (I + \rho_l W_l + \rho_l^2 W_l^2 + \rho_l^3 W_l^3 + \dots) Q_h \\
&= (Q_h + \rho_l W_l Q_h + \rho_l^2 W_l^2 Q_h + \rho_l^3 W_l^3 Q_h + \dots) \\
&= Q_h (I + \rho_l W_l + \rho_l^2 W_l^2 + \rho_l^3 W_l^3 + \dots) = Q_h (I - \rho_l W_l)^{-1} \\
&= Q_h \Lambda_l
\end{aligned}
$$

3 Single Equation Estimation

Without loss of generality, let us consider the first equation of the system

$$
y_1 = Z_1 \delta_1 + \upsilon_1 = Z_1 \delta_1 + \Lambda_1 \epsilon_1 \tag{8}
$$

To get spatial within and spatial between 2SLS estimators, we need a procedure which both resolves the endogeneity problem and spatial correlation. Thus, one need to combine instrumental variables (IV) and spatial generalized least squares (GLS) estimator. Hence, applying transformation Q_h on equation (8) gives

$$
y_1^{(h)} = Z_1^{(h)} \delta_1 + \Lambda_1 \epsilon_1^{(h)} \tag{9}
$$

where $y_1^{(h)} = Q_h y_1$, $Z_1^{(h)} = Q Z_1$ and $\epsilon_1^{(h)} = Q_h \epsilon_1$ with,

$$
\mathrm{E}(\Lambda \epsilon_1^{(h)}) = \Lambda_1 Q_h \mathrm{E}(\epsilon_1) = 0
$$

and using equation (7),

$$
\mathrm{var}\left(\Lambda_1 \epsilon_1^{(h)}\right) = \mathrm{var}\left(Q_h \upsilon_1\right) = \sigma_{h_{11}}^2 Q_h \Lambda_{11}
$$

applying Aitken procedure on the following equation

$$
X_h' y_1^{(h)} = X_h' Z_1^{(h)} \delta_1 + X_h' \Lambda_1 \epsilon_1^{(h)} \tag{10}
$$

with $X_h = Q_h X$ and

$$
\mathrm{var}\left(X_h' \Lambda_1 \epsilon_1^{(h)}\right) = \sigma_{h_{11}}^2 X_h' \Lambda_{11} X_h \tag{11}
$$

gives

$$
\begin{aligned}
\hat{\delta}^{(h)}_{1,S2SLS} &= \left[Z_1^{(h)\prime} X_h \left(\sigma^2_{h_{11}} X_h^\prime \Lambda_{11} X_h\right)^{-1} X_h^\prime Z_1^{(h)}\right]^{-1} \left[Z_1^{(h)\prime} X_h \left(\sigma^2_{h_{11}} X_h^\prime \Lambda_{11} X_h\right)^{-1} X_h^\prime y_1^{(h)}\right] \\
&= \left(Z_1^{(h)\prime} P_h Z_1^{(h)}\right)^{-1} \left(Z_1^{(h)\prime} P_h y_1^{(h)}\right)
\end{aligned} \tag{12}
$$

where $P_h = X_h \left(X_h^\prime \Lambda_{11} X_h\right)^{-1} X_h^\prime$ and with a known ρ_1, we estimate the variance components by

$$
\hat{\sigma}^2_{h_{11}} = \frac{\left(y_1^\star - Z_1^\star \hat{\delta}^{(h)}_{1,S2SLS}\right)^\prime Q_h \left(y_1^\star - Z_1^\star \hat{\delta}^{(h)}_{1,S2SLS}\right)}{\mathrm{tr}(Q_h)} \tag{13}
$$

where $y_1^\star = \Lambda^{-1} y_1$ and $Z_1^\star = \Lambda^{-1} Z_1$. In equation (13), the adjustment on y_1 and Z_1 is important to remove the spatial effect on residuals while the matrix Q_h removes the corresponding specific effect.
The important thing to notice in this case is that the unknown variance components $\sigma^2_{h_{11}}$ are estimated using classical one way 2SLS or 3SLS residuals and not OLS or general 2SLS residuals of the transformed system (9). This is due to the simultaneous nature of (9).
With a known ρ_1, equation (12) gives the spatial *within* and the spatial *between* 2SLS estimators of δ_1 for $h = 0$ and $h = 1$ respectively. Note that, if $\rho_1 = 0$ this implies that $\Lambda_{11} = I_{TN}$ and the projection matrix P_h simply P_{X_h}. This means that when ρ_1 is null, we fall in the classical simultaneous panel data models.
An estimate of the asymptotic covariance matrix of $\hat{\delta}_{1,2SLS}$ is obtained by multiply $\sigma^2_{h_{11}}$ by the inverted matrix on the righ-hand side of equation (12) i.e.

$$
\mathrm{var}\left(\hat{\delta}_{1,S2SLS}\right) = \sigma^2_{h_{11}} \left(Z_1^{(h)\prime} P_h Z_1^{(h)}\right)^{-1} \tag{14}
$$

Iterating equation (10), stacking these two transformed equations as a system and noting that δ_1 is the same for these two transformed equations, we can get a more efficient estimator of δ_1, by applying an Aitken estimation procedure to the following system

$$
\begin{pmatrix} X_0^\prime y_1^{(0)} \\ X_1^\prime y_1^{(1)} \end{pmatrix} = \begin{pmatrix} X_0^\prime Z_1^{(0)} \\ X_1^\prime Z_1^{(1)} \end{pmatrix} \delta_1 + \begin{pmatrix} X_0^\prime \upsilon_1^{(0)} \\ X_1^\prime \upsilon_1^{(1)} \end{pmatrix} \tag{15}
$$

Proposition 1. In a simultaneous equation with spatially autocorrelated error components, if the component ρ_1 is known, the spatial error component two-stage least squares, says $\hat{\delta}_{1,SEC2SLS}$, is:

$$
\hat{\delta}_{1,SEC2SLS} = \left(\frac{Z_1^{\prime(0)} P_0 Z_1^{(0)}}{\hat{\sigma}^2_{0_{11}}} + \frac{Z_1^{\prime(1)} P_1 Z_1^{(1)}}{\hat{\sigma}^2_{1_{11}}}\right)^{-1} \left(\frac{Z_1^{\prime(0)} P_0 y_1^{(0)}}{\hat{\sigma}^2_{0_{11}}} + \frac{Z_1^{\prime(1)} P_1 y_1^{(1)}}{\hat{\sigma}^2_{1_{11}}}\right) \tag{16}
$$

where, $P_h = X_h \left(X_h'\Lambda_{11}X_h\right)^{-1} X_h'$, $X_h = Q_hX$, $y_1^{(h)} = Q_hy_1$, $Z_1^{(h)} = QZ_1$ for $h = 0$ and $h = 1$. With $\hat{\sigma}^2_{0_{11}}$ and $\hat{\sigma}^2_{1_{11}}$ defined in equation (13).

Proof. Since

$$\text{var}\begin{pmatrix} X_0'\upsilon_1^{(0)} \\ X_1'\upsilon_1^{(1)} \end{pmatrix} = \text{diag}\left(\sigma^2_{h_{lq}} X_h'\Lambda_{11}X_h\right) \tag{17}$$

we simply apply GLS on system (15). □

Practically, the component ρ_l is unknow then we need a procedure to derive a consistent estimate of ρ_l. A solution can be obtained from Kapoor et al. (2007) GMM procedure for SAR and from Fingleton (2008) for SMA.

4 Estimation of the component ρ_l

For notation convenience, let

$$\overset{a}{\bar{\epsilon}} = W^a\epsilon,$$

this means that $\overset{0}{\bar{\epsilon}} = W^0\epsilon = \epsilon$, $\overset{1}{\bar{\epsilon}} = W\epsilon$ and $\overset{2}{\bar{\epsilon}} = W^2\epsilon = \bar{\bar{\epsilon}}$. Such that

$$Q_h\text{E}(\epsilon_1\epsilon_1') = Q_h\Omega_{\epsilon_{11}} = \sigma^2_{h_{11}}Q_h = \begin{cases} \sigma^2_{0_{11}}Q_0 & \text{for} \quad h = 0 \\ \sigma^2_{1_{11}}Q_1 & \text{for} \quad h = 1 \end{cases} \tag{18}$$

and using the following general relation:

$$\begin{aligned} \text{E}\left(\overset{a}{\bar{\epsilon}}{}_1' Q_h \overset{b}{\bar{\epsilon}}_1\right) &= \text{E}\left(\epsilon_1' \underbrace{W_1' \ldots W_1'}_{a \text{ times}} Q_h \underbrace{W_1 \ldots W_1}_{b \text{ times}} \epsilon_1\right) \\ &= \text{E}\left(\epsilon_1'(W_1')^a Q_h W_1^b \epsilon_1\right) \\ &= \text{tr}\left((W_1')^a W_1^b Q_h \Omega_{\epsilon_{11}}\right) \\ &= \sigma^2_{h_{11}} \text{tr}(B_h) \text{tr}\left((W_{1N}')^a W_{1N}^b\right) \end{aligned} \tag{19}$$

with $\text{tr}(W_{1N}^0) = \text{tr}(I_N) = N$ and $\text{tr}(W_{1N}) = 0$; the general form of the six moment conditions can be derived as follow:

$$\mathrm{E}(\epsilon_1' Q_h \epsilon_1) = \sigma^2_{h_{11}} \mathrm{tr}\,(Q_h) \tag{20}$$
$$\mathrm{E}(\bar{\epsilon}_1' Q_h \bar{\epsilon}_1) = \sigma^2_{h_{11}} \mathrm{tr}\,(B_h)\,\mathrm{tr}\,(W_{1N} W_{1N}') \tag{21}$$
$$\mathrm{E}(\bar{\epsilon}_1' Q_h \epsilon_1) = \sigma^2_{h_{11}} \mathrm{tr}\,(B_h)\,\mathrm{tr}\,(W_{1N}) = 0 \tag{22}$$

which lead to general system for $T \geq 2$

$$\mathrm{E}\begin{bmatrix} \epsilon_1' Q_h \epsilon_1/\mathrm{tr}(Q_h) \\ \bar{\epsilon}_1' Q_h \bar{\epsilon}_1/\mathrm{tr}(Q_h) \\ \bar{\epsilon}_1' Q_h \epsilon_1/\mathrm{tr}(Q_h) \end{bmatrix} = \sigma^2_{h_{11}} \begin{bmatrix} 1 \\ \mathrm{tr}\,(W_{1N} W_{1N}')/N \\ 0 \end{bmatrix} \tag{23}$$

To get each equation, one should replace each notation by its corresponding form (see Kapoor et al., 2007, for more details). Our three GMM estimators of ρ_1, $\sigma^2_{0_{11}}$ and $\sigma^2_{1_{11}}$ are based on these moment relationships. If ϵ_1 were observed, then $\epsilon_1' Q_h \epsilon_1/\mathrm{tr}(Q_h)$ represents the (unbiased) analysis of variance estimators of $\sigma^2_{h_{11}}$.

4.1 SAR Process

For spatial autoregressive process,

$$\upsilon_1 = (I - \rho_1 W_1)^{-1} \epsilon_1 \implies \begin{cases} \epsilon_1 = \upsilon_1 - \rho_1 \bar{\upsilon}_1 \\ \bar{\epsilon}_1 = \bar{\upsilon}_1 - \rho_1 \bar{\bar{\upsilon}}_1 \end{cases}$$

Substituting these expressions for ϵ_1 and $\bar{\epsilon}_1$ into equations (20) to (22), we obtain the general form of system of three equations involving the second moments of υ_1, $\bar{\upsilon}_1$ and $\bar{\bar{\upsilon}}_1$.

$$\begin{cases} \epsilon_1' Q_h \epsilon_1 &= \upsilon_1' Q_h \upsilon_1 - \rho_1 \left(\upsilon_1' Q_h \bar{\upsilon}_1 + \bar{\upsilon}_1' Q_h \upsilon_1\right) + \rho_1^2 \bar{\upsilon}_1' Q_h \bar{\upsilon}_1 \\ \bar{\epsilon}_1' Q_h \bar{\epsilon}_1 &= \bar{\upsilon}_1' Q_h \bar{\upsilon}_1 - \rho_1 \left(\bar{\upsilon}_1' Q_h \bar{\bar{\upsilon}}_1 + \bar{\bar{\upsilon}}_1' Q_h \bar{\upsilon}_1\right) + \rho_1^2 \bar{\bar{\upsilon}}_1' Q_h \bar{\bar{\upsilon}}_1 \\ \bar{\epsilon}_1' Q_h \epsilon_1 &= \bar{\upsilon}_1' Q_h \upsilon_1 - \rho_1 \left(\bar{\upsilon}_1' Q_h \bar{\upsilon}_1 + \bar{\bar{\upsilon}}_1' Q_h \upsilon_1\right) + \rho_1^2 \bar{\bar{\upsilon}}_1' Q_h \bar{\upsilon}_1 \end{cases} \tag{24}$$

Hence the six equations can be easily obtained by simply iterate $h = 0,\ 1$ on system (24). This system involves ρ_1 and $\sigma^2_{h_{11}}$ and can be expressed as:

$$\Gamma^{1h} \begin{pmatrix} \rho_1 \\ \rho_1^2 \\ \sigma^2_{h_{11}} \end{pmatrix} - \Theta^{1h} = 0 \tag{25}$$

where

$$\Gamma^{1h} = \begin{bmatrix} 2\gamma_1^{1h} & -\gamma_2^{1h} & 1 \\ 2\gamma_3^{1h} & -\gamma_4^{1h} & \gamma_5^{1h} \\ \left(\gamma_2^{1h} + \gamma_6^{1h}\right) & -\gamma_3^{1h} & 0 \end{bmatrix} \quad \Theta^{1h} = \begin{bmatrix} \gamma_0^{1h} \\ \gamma_2^{1h} \\ \gamma_1^{1h} \end{bmatrix}$$

the elements γ_i^{1h} have the following form,

$$\begin{aligned} \gamma_1^{1h} &= \frac{(\bar{\upsilon}_1' Q_h \upsilon_1)}{\text{tr}\,(Q_h)}, \quad \gamma_2^{1h} = \frac{(\bar{\upsilon}_1' Q_h \bar{\upsilon}_1)}{\text{tr}\,(Q_h)}, \quad \gamma_3^{1h} = \frac{(\bar{\upsilon}_1' Q_h \bar{\bar{\upsilon}}_1)}{\text{tr}\,(Q_h)} \\ \gamma_4^{1h} &= \frac{(\bar{\bar{\upsilon}}_1' Q_h \bar{\bar{\upsilon}}_1)}{\text{tr}\,(Q_h)}, \quad \gamma_5^{1h} = \frac{\text{tr}(W_{1N}' W_{1N})}{N}, \quad \gamma_6^{1h} = \frac{(\bar{\bar{\upsilon}}_1' Q_h \upsilon_1)}{\text{tr}\,(Q_h)}, \quad \gamma_0^{1h} = \frac{(\upsilon_1' Q_h \upsilon_1)}{\text{tr}\,(Q_h)}. \end{aligned}$$

4.2 SMA Process

For spatial moving average process,

$$\upsilon_1 = (I + \rho_1 W_1)\epsilon_1 \implies \begin{cases} \upsilon_1 = \epsilon_1 + \rho_1 \bar{\epsilon}_1 \\ \bar{\upsilon}_1 = \bar{\epsilon}_1 + \rho_1 \bar{\bar{\epsilon}}_1 \end{cases}$$

then,

$$\begin{aligned} \upsilon_1' Q_h \upsilon_1 &= \epsilon_1' Q_h \epsilon_1 + \rho_1 \left(\epsilon_1' Q_h \bar{\epsilon}_1 + \bar{\epsilon}_1' Q_h \epsilon_1\right) + \rho_1^2 \bar{\epsilon}_1' Q_h \bar{\epsilon}_1 \\ \bar{\upsilon}_1' Q_h \bar{\upsilon}_1 &= \bar{\epsilon}_1' Q_h \bar{\epsilon}_1 + \rho_1 \left(\bar{\epsilon}_1' Q_h \bar{\bar{\epsilon}}_1 + \bar{\bar{\epsilon}}_1' Q_h \bar{\epsilon}_1\right) + \rho_1^2 \bar{\bar{\epsilon}}_1' Q_h \bar{\bar{\epsilon}}_1 \\ \bar{\upsilon}_1' Q_h \upsilon_1 &= \bar{\epsilon}_1' Q_h \epsilon_1 + \rho_1 \left(\bar{\epsilon}_1' Q_h \bar{\epsilon}_1 + \bar{\bar{\epsilon}}_1' Q_h \epsilon_1\right) + \rho_1^2 \bar{\bar{\epsilon}}_1' Q_h \bar{\epsilon}_1 \end{aligned} \tag{26}$$

to obtain the expectation of each equation of system (26), we use relation (19)

$$\begin{aligned} \text{E}\left(\upsilon_1' Q_h \upsilon_1\right) &= \sigma_{h_{11}}^2 \text{tr}(B_h) \left\{ N + \rho_1 \left[\text{tr}\,(W_{1N}) + \text{tr}\,(W_{1N})\right] + \rho_1^2 \text{tr}\,(W_{1N}' W_{1N}) \right\} \\ \text{E}\left(\bar{\upsilon}_1' Q_h \bar{\upsilon}_1\right) &= \sigma_{h_{11}}^2 \text{tr}(B_h) \left\{ \text{tr}\,(W_{1N}' W_{1N}) + \rho_1 \left[\text{tr}\left(W_{1N}' W_{1N}^2\right) + \text{tr}\left((W_{1N}')^2 W_{1N}\right)\right] + \rho_1^2 \text{tr}\left((W_{1N}')^2 W_{1N}^2\right) \right\} \\ \text{E}\left(\bar{\upsilon}_1' Q_h \upsilon_1\right) &= \sigma_{h_{11}}^2 \text{tr}(B_h) \left\{ \text{tr}\,(W_{1N}) + \rho_1 \left[\text{tr}\,(W_{1N}' W_{1N}) + \text{tr}\left(W_{1N}^2\right)\right] + \rho_1^2 \text{tr}\left((W_{1N}')^2 W_{1N}\right) \right\} \end{aligned}$$

which lead to

$$\mathrm{E}\left(\frac{\upsilon_1' Q_h \upsilon_1}{\mathrm{tr}\,(Q_h)}\right) = \sigma^2_{h_{11}}\left[1 + \rho_1^2 \frac{\mathrm{tr}\,(W_{1N}' W_{1N})}{N}\right] \tag{27}$$

$$\mathrm{E}\left(\frac{\bar{\upsilon}_1' Q_h \bar{\upsilon}_1}{\mathrm{tr}\,(Q_h)}\right) = \sigma^2_{h_{11}}\left[\frac{\mathrm{tr}\,(W_{1N}' W_{1N})}{N} + 2\rho_1 \frac{\mathrm{tr}\,(W_{1N}' W_{1N}^2)}{N} + \rho_1^2 \frac{\mathrm{tr}\,((W_{1N}')^2 W_{1N}^2)}{N}\right] \tag{28}$$

$$\mathrm{E}\left(\frac{\bar{\upsilon}_1' Q_h \upsilon_1}{\mathrm{tr}\,(Q_h)}\right) = \sigma^2_{h_{11}}\rho_1\left[\frac{\mathrm{tr}\,(W_{1N}' W_{1N})}{N} + \frac{\mathrm{tr}\,(W_{1N}^2)}{N} + \rho_1 \frac{\mathrm{tr}\,((W_{1N}')^2 W_{1N})}{N}\right] \tag{29}$$

Ignoring the expectations, and put these equations together using the 3×3 matrix Γ^{1h}, the 3×1 vector Θ^{1h},

$$\gamma_1^1 = \frac{\mathrm{tr}\,(W_{1N}' W_{1N})}{N}, \gamma_2^1 = \frac{\mathrm{tr}\,(W_{1N}' W_{1N}^2)}{N}, \gamma_3^1 = \frac{\mathrm{tr}\,(W_{1N}^2)}{N}, \gamma_4^1 = \frac{\mathrm{tr}\,((W_{1N}')^2 W_{1N}^2)}{N}$$

so that

$$\Gamma^{1h}\begin{pmatrix}\sigma^2_{h_{11}}\\ \rho_1\sigma^2_{h_{11}}\\ \rho_1^2\sigma^2_{h_{11}}\end{pmatrix} - \Theta^{1h} = 0 \tag{30}$$

where

$$\Gamma^{1h} = \begin{bmatrix}1 & 0 & \gamma_1^1\\ \gamma_1^1 & 2\gamma_2^1 & \gamma_4^1\\ 0 & (\gamma_1^1 + \gamma_3^1) & \gamma_2^1\end{bmatrix}, \quad \Theta^{1h} = \frac{1}{\mathrm{tr}\,(Q_h)}\begin{bmatrix}\upsilon_1' Q_h \upsilon_1\\ \bar{\upsilon}_1' Q_h \bar{\upsilon}_1\\ \bar{\upsilon}_1' Q_h \upsilon_1\end{bmatrix}$$

4.3 Estimation

The estimation procedure of SAR (respectively SMA) process comprises two (02) stages. At stage one, because of simultaneity problem, we obtain IV estimates of $\delta_1^\star$ and hence residuals $\upsilon_1^\star = y_1 - Z_1\delta_1^\star$; At stage two, we use these IV residuals to obtain the estimates g^{1h} of Θ^{1h} and G^{1h} of Γ^{1h}. Then a sample analogue to equation (25) or (30) in terms of $\upsilon_1^\star$, $\bar{\upsilon}_1^\star$ and $\bar{\bar{\upsilon}}_1^\star$ is

$$G^{1h}\psi^h - g^{1h} = e(\rho_1, \sigma^2_{h_{11}}) \tag{31}$$

in which

$$\psi^h = \begin{cases}\psi^h_{SAR} = \begin{pmatrix}\rho_1 & \rho_1^2 & \sigma^2_{h_{11}}\end{pmatrix}' & \text{for SAR}\\ \psi^h_{SMA} = \begin{pmatrix}\sigma^2_{h_{11}} & \rho_1\sigma^2_{h_{11}} & \rho_1^2\sigma^2_{h_{11}}\end{pmatrix}' & \text{for SMA}\end{cases}$$

and $e(\rho_1, \sigma^2_{h_{11}})$ is a vector of residuals.

The GMM estimators of ρ_1 and $\sigma^2_{h_{11}}$ are the solution of the sample counterpart of the six equations given above. Following Kapoor et al. (2007) and using their results in the context

of simultaneous panel data, we only use the first three moments which do not involve $\sigma^2_{1_{11}}$ and yield estimates of ρ_1 and $\sigma^2_{0_{11}}$. The fourth moment condition is then used to solve for $\sigma^2_{1_{11}}$ given estimates of ρ_1 and ρ_1. Since Fingleton (2008) extended this GMM estimator to the SMA panel data model with random effects, we also use and adapte his results for the SMA case.
The non-linear least square estimator is therefore given by

$$(\tilde{\rho}_1, \tilde{\sigma}^2_{h_{11}}) = \arg\min\{e(\rho_1, \sigma^2_{h_{11}})' e(\rho_1, \sigma^2_{h_{11}})\} \tag{32}$$

Proposition 2. In a simultaneous equation with spatially autocorrelated error components and an unknow ρ_1 the spatial error component two-stage least square, says $\hat{\delta}_{1,SEC2SLS}$, is:

$$\hat{\delta}_{1,SEC2SLS} = \left(\frac{Z_1'^{(0)} P_0 Z_1^{(0)}}{\hat{\sigma}^2_{0_{11}}} + \frac{Z_1'^{(1)} P_1 Z_1^{(1)}}{\hat{\sigma}^2_{1_{11}}} \right)^{-1} \left(\frac{Z_1'^{(0)} P_0 y_1^{(0)}}{\hat{\sigma}^2_{0_{11}}} + \frac{Z_1'^{(1)} P_1 y_1^{(1)}}{\hat{\sigma}^2_{1_{11}}} \right) \tag{33}$$

where $\hat{\rho}_1$, $\sigma^2_{\xi_{11}}$ and $\sigma^2_{1_{11}}$ are directly obtained from non-linear least square estimators. $P_h = X_h \left(X_h' \Lambda_{11} X_h\right)^{-1} X_h'$, $X_h = Q_h X$, $y_1^{(h)} = Q_h y_1$, $Z_1^{(h)} = Q Z_1$ for $h = 0$ and $h = 1$.

Proof. We replace ρ_1, $\sigma^2_{\xi_{11}}$ and $\sigma^2_{1_{11}}$ by their consistent estimate $\hat{\rho}_1$, $\hat{\sigma}^2_{\xi_{11}}$ and $\hat{\sigma}^2_{1_{11}}$ in equation (32). □

5 System Estimation

The single equation estimation described in the previous section provides consistent estimators in the presence of a spatially autocorrelated error. But like all single equation estimators, this SEC2SLS estimator ignores the cross equation correlation between υ_l and υ_q, and the information content of the full system of simultaneous equations. This is bound to result in loss of efficiency.
This section focuses on 3SLS system estimation which utilizes the correlation across equations and should lead to gains in efficiency over its 2SLS counterpart. Of course, this system estimation has to handle the spatial autocorrelation structure, the presence of right hand-side endogenous variables as well as individual random effect.

The system of L equations can be obtained from equation (1), by iterate $l = 1, \ldots, L$ and stack these L equations

$$y = Z\delta + \upsilon \tag{34}$$

where, $y' = \begin{pmatrix} y_1' & \cdots & y_L' \end{pmatrix}$, $Z = \text{diag}(Z_l)$, $\delta' = \begin{pmatrix} \delta_1' & \cdots & \delta_L' \end{pmatrix}$ and $\upsilon' = \begin{pmatrix} \upsilon_1' & \cdots & \upsilon_L' \end{pmatrix}$ with $Z_l = \begin{bmatrix} Y_l & X_l \end{bmatrix}$. The disturbance process of the system can be written as:

$$\upsilon = \Lambda\epsilon = \begin{cases} \text{diag}(\rho_l W_l)\upsilon + \epsilon = \text{diag}(I - \rho_l W_l)^{-1}\epsilon & \text{for SAR} \\ \text{diag}(\rho_l W_l)\epsilon + \epsilon = \text{diag}(I + \rho_l W_l)\epsilon & \text{for SMA} \end{cases} \tag{35}$$

where $\Lambda = \text{diag}\left(\Lambda_l\right)$ and $\epsilon' = \begin{pmatrix}\epsilon'_1 & \cdots & \epsilon'_L\end{pmatrix}$; with

$$\epsilon = (I_L \otimes \iota_T \otimes I_N)\eta + \xi = I_L \otimes Z_\eta \eta + \xi \tag{36}$$

where $\eta' = \begin{pmatrix}\eta'_1 & \cdots & \eta'_L\end{pmatrix}$ and $\xi' = \begin{pmatrix}\xi'_1 & \cdots & \xi'_L\end{pmatrix}$.
We note that if the row-standardized spatial matrix W_l is common to all equation,i.e.,$W_l = W$ hence,

$$\begin{pmatrix} \rho_1 W & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \rho_L W \end{pmatrix} = \text{diag}(\rho_l) \otimes W = \rho \otimes W \tag{37}$$

The covariance matrix of innovations for the system Ω_ϵ is

$$\Omega_\epsilon = \left(\Omega_{\epsilon_{lq}}\right) = \left(\sigma^2_{0_{lq}}\right) \otimes Q_0 + \left(\sigma^2_{1_{lq}}\right) \otimes Q_1 = \Sigma_0 \otimes Q_0 + \Sigma_1 \otimes Q_1 \tag{38}$$

with $\Sigma_1 = \left(\sigma^2_{1_{lq}}\right)$ and $\Sigma_0 = \left(\sigma^2_{0_{lq}}\right)$. Using equation (35) the covariance matrix of the disturbance υ can be written as follow:

$$\Omega_\upsilon = \Lambda\Omega_\epsilon\Lambda' = \Lambda\left(\Sigma_0 \otimes Q_0\right)\Lambda' + \Lambda\left(\Sigma_1 \otimes Q_1\right)\Lambda'$$

however,

$$\begin{aligned} \Lambda\left(\Sigma_h \otimes Q_h\right)\Lambda' &= \text{diag}(\Lambda_l)\left(\sigma^2_{h_{lq}} Q_h\right)\text{diag}(\Lambda'_l) \\ &= \left(\sigma^2_{h_{lq}} \Lambda_l Q_h \Lambda'_q\right) \\ &= \left(\sigma^2_{h_{lq}} Q_h \Lambda_{lq}\right) \end{aligned}$$

this implies that

$$\Omega_\upsilon = \left(I_L \otimes Q_0\right)\Delta_{\Sigma_0} + \left(I_L \otimes Q_1\right)\Delta_{\Sigma_1} \tag{39}$$

where $\Delta_{\Sigma_h} = \left(\sigma^2_{h_{lq}} \Lambda_{lq}\right)$. Starting with equation (34) and applying the transformation $I_L \otimes Q_h$ we get:

$$y^{(h)} = Z^{(h)}\delta + \Lambda\epsilon^{(h)} \tag{40}$$

where $y^{(h)} = (I_L \otimes Q_h)y$, $Z^{(h)} = (I_L \otimes Q_h)Z$ and $\epsilon^{(h)} = (I_L \otimes Q_h)\epsilon$ with,

$$\mathrm{E}(\Lambda\epsilon^{(h)}) = \Lambda(I_L \otimes Q_h)\mathrm{E}(\epsilon) = 0$$

and,

$$\mathrm{var}\left(\Lambda\epsilon^{(h)}\right) = (I_L \otimes Q_h)\,\Delta_{\Sigma_h} \tag{41}$$

Premultiply equation (40) by $x_h = I_L \otimes X_h$ gives

$$x_h'y^{(h)} = x_h'Z^{(h)}\delta + x_h'\Lambda\epsilon^{(h)} \tag{42}$$

with

$$\mathrm{var}\left(x_h'\Lambda\epsilon^{(h)}\right) = x_h'\mathrm{var}\left(\Lambda\epsilon^{(h)}\right)x_h = x_h'\Delta_{\Sigma_h}x_h \tag{43}$$

where $\Delta_{\Sigma_h} = \left(\sigma^2_{h_{lq}}\Lambda_{lq}\right)$. Applying Aitken procedure on equation (42) give us

$$\begin{aligned}\hat{\delta}^{(h)}_{S3SLS} &= \left[Z^{(h)\prime}x_h\left(x_h'\Delta_{\Sigma_h}x_h\right)^{-1}x_h'Z^{(h)}\right]^{-1}\left[Z^{(h)\prime}x_h\left(x_h'\Delta_{\Sigma_h}x_h\right)^{-1}x_h'y^{(h)}\right]\\ &= \left(Z^{(h)\prime}P_hZ^{(h)}\right)^{-1}\left(Z^{(h)\prime}P_hy^{(h)}\right)\end{aligned} \tag{44}$$

with $P_h = x_h\left(x_h'\Delta_{\Sigma_h}x_h\right)^{-1}x_h'$.
Equation (44) gives the spatial *within* and the spatial *between* 3SLS estimator of δ for $h = 0$ and 1 respectively. An estimate of the asymptotic covariance matrix of $\hat{\delta}^{(h)}_{S3SLS}$ is given by the inverted matrix on the righ-hand side of equation (44).

Iterate equation (42), stack these two transformed equations as a system and noting that δ is the same for these two transformed equations, we can get a more efficient estimator of δ. This is done by applying an Aitken estimation procedure to the following system:

$$\begin{pmatrix} x_0'y^{(0)} \\ x'y^{(1)} \end{pmatrix} = \begin{pmatrix} x_0'Z^{(0)} \\ x'Z^{(1)} \end{pmatrix}\delta + \begin{pmatrix} x_0'\upsilon^{(0)} \\ x'\upsilon^{(1)} \end{pmatrix} \tag{45}$$

Proposition 3. In a simultaneous equation with spatially autocorrelated error components, the spatial error component three stage least squares, says $\hat{\delta}_{SEC3SLS}$, is:

$$\hat{\delta}_{SEC3SLS} = \left(Z'^{(0)}P_0Z^{(0)} + Z_1'^{(1)}P_1Z_1^{(1)}\right)^{-1}\left(Z'^{(0)}P_0y^{(0)} + Z_1'^{(1)}P_1y_1^{(1)}\right) \tag{46}$$

where, $P_h = x_h\left(x_h'\Delta_{\Sigma_h}x_h\right)^{-1}x_h'$; $y^{(h)} = I_M \otimes Q_hy$, $Z^{(h)} = I_M \otimes Q_hZ$ for $h = 0$ and $h = 1$. With

$$\hat{\sigma}^2_{h_{lq}} = \frac{\left(y_l^\star - Z_l^\star\hat{\delta}^{(h)}_{l,S2SLS}\right)'Q_h\left(y_q^\star - Z_q^\star\hat{\delta}^{(h)}_{q,S2SLS}\right)}{\mathrm{tr}(Q_h)} \tag{47}$$

Proof. Straightforward. □

6 Monte Carlo Investigation

6.1 Design of sampling

The purposes of our Monte Carlo experiment are threefolds: Firstly, we study the small sample behavior of our proposed estimators that can handle endogeneity, spatial error correlation and random individual effects in function of spatial coefficient, spatial matrix, variance covariance of specific component and the increase of time. These estimators are compared with those that may ignore one or more of these symptoms. For example, OLS ignores all these symptoms, while EC-2SLS only ignore spatial error correlation. Secondly, we also investigate the gain in efficiency; for example when we move from usual one way to spatial one way estimator. Also, when we move from spatial two stage that does not take into account simultaneity, to spatial three stage least squares. Thirdly, we study the sample properties of the spatial component ρ_l, which is necessary to get $\sigma^2_{h_{lq}}$ and our spatial estimators. We note that, the estimations of ρ_l, $\sigma^2_{h_{ll}}$ are done on each equation l; and we use four (04) values of ρ_l namely, -0.8, -0.4, 0.4 and 0.8. Hence, we can write the linear simultaneous equation model in equation (1) as:

$$\Gamma y_{it} + \Lambda x_{it} = \upsilon_{it} \tag{48}$$

Here y_{it}, x_{it} and υ_{it} are column vectors of dimensions 2, 4 and 2 respectively. We simplify the Monte Carlo design by using the same weight matrix W in both equations. The disturbance υ_{itl}, $l = 1, 2$, for each equation has the following form

$$\upsilon_{itl} = \begin{cases} (I - \rho_lW)^{-1}\epsilon_l & \text{for SAR} \\ (I + \rho_lW)\epsilon_l & \text{for SMA} \end{cases}$$

where $\epsilon_{itl} = \eta_{il} + \xi_{itl}$. Γ is a 2×2 matrix of coefficients of current endogenous variables and Λ is a 2×4 matrix coefficients of predetermined variables.

$\Gamma = \begin{pmatrix} 1 & 0.5 \\ 4 & 1 \end{pmatrix}$, $\Lambda = \begin{pmatrix} 2 & -1.5 & 0 & 0 \\ 0 & 0 & 3 & -1.8 \end{pmatrix}$

There are four exogenous variables X_{11}, X_{12}, X_{21} and X_{22} in the system, two for each equation. The data generating process for the exogenous variables follows the approach used in Baltagi et al. (2013)

$$x_{p,it} = \zeta_{p,i} + z_{p,it} \quad p = 11, 12, 21, 22$$

where $\zeta_{p,i} \rightsquigarrow iidU[-10, 10]$, and $z_{p,it} \rightsquigarrow iidU[-5, 5]$. We follow two steps to generate the error terms. First, we generate $2(N + NT)$ independent $\mathcal{N}(0, 1)$ random numbers. For each equation, the first $2N$ are used for generating first cross section specific effects and the remaining $2NT$ are used to generate the idiosyncratic errors.
Second, we transform these $\mathcal{N}(0, 1)$ disturbances to obtain the appropriate covariance matrices Ω_η, Ω_ξ respectively. Four combinations are considered:

V1. $\Omega_\eta = \begin{pmatrix} 16 & 8 \\ 8 & 16 \end{pmatrix}$ and $\Omega_\xi = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$

V2. $\Omega_\eta = \begin{pmatrix} 12 & 6 \\ 6 & 12 \end{pmatrix}$ and $\Omega_\xi = \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}$

V3. $\Omega_\eta = \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}$ and $\Omega_\xi = \begin{pmatrix} 12 & 6 \\ 6 & 12 \end{pmatrix}$

V4. $\Omega_\eta = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$ and $\Omega_\xi = \begin{pmatrix} 16 & 8 \\ 8 & 16 \end{pmatrix}$

For all experiments, we keep the total variance fixed at $\Omega_\epsilon = \begin{pmatrix} 20 & 10 \\ 10 & 20 \end{pmatrix}$

For the spatial weights matrices, we use regular[4] structures. We decide to use four weight matrices, W_1,W_3,W_7 and W_9, which essentially differ in their degree of sparseness (see Figure 7 in the Appendix). In fact, the matrix W_J where J is a positive integer is labelled as "J ahead and J behind". Since in panel data many studies are not done in all the countries, we relax the hypothesis of a circular world in the construction of the matrix W_J. We consider several individuals $N = 25$ and time dimensions $T = (7, 10, 15)$. First, we consider five simultaneous equation estimators of the one-way error component model which ignore spatial dependence:

1. Ordinary Least Squares (OLS).

2. Two Stage Least Square (2SLS).

3. Fixed Effects Two Stage Least Squares (FE-2SLS).

4. Error Component Two Stage Least Squares (EC-2SLS).

5. Error Component Three Stage Least Squares (EC-3SLS).

[4] Irregular lattices structures are left for application on real data.

Second, we consider our two simultaneous equation estimators which take into account cross-section spatial dependence:

1. Spatial Error component Two Stage Least Squares (SEC-2SLS).

2. Spatial Error Component Three Stage Least Squares (SEC-3SLS).

To sum up, we will have a total of sixteen (16) Monte Carlo designs (see Table 1). We run 1000 replications for each experiment.

Table 1: Monte Carlo designs

No	$\sigma^2_{\eta\iota q}$	$\sigma^2_{\xi\iota q}$	W_N	T	ρ_1	ρ_2	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}
Covariance												
1	16	4	W_3	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
2	12	8	W_3	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
3	8	12	W_3	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
4	4	16	W_3	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
Spatial matrix												
5	12	8	W_1	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
6	12	8	W_3	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
7	12	8	W_7	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
8	12	8	W_9	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
Time												
9	12	8	W_3	7	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
10	12	8	W_3	10	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
11	12	8	W_3	15	-0.8	-0.3	-0.5	-2	1.5	-4	-3	1.8
Spatial coefficient												
12	12	8	W_3	7	-0.8	-0.8	-0.5	-2	1.5	-4	-3	1.8
13	12	8	W_3	7	-0.4	-0.4	-0.5	-2	1.5	-4	-3	1.8
14	12	8	W_3	7	0.0	0.0	-0.5	-2	1.5	-4	-3	1.8
15	12	8	W_3	7	0.4	0.4	-0.5	-2	1.5	-4	-3	1.8
16	12	8	W_3	7	0.8	0.8	-0.5	-2	1.5	-4	-3	1.8

6.2 Efficiency Criteria

To compare the performance of these estimators, we use three criteria. The first is an adjusted version of the root mean square error (RMSE) criterion proposed by Kelejian and Prucha (1998):

$$RMSE^*(\hat{\alpha}_k) = \left[bias^2(\hat{\alpha}_k) + \left(\frac{IQ(\hat{\alpha}_k)}{1.35} \right)^2 \right]^{1/2}$$

where median is used instead of mean for *bias*. So bias is the difference between median and the true parameter. IQ is the inter-quartile range defined as the difference between the 0.75 quantile and the 0.25 quantile of our estimates, $\hat{\alpha}_k$ is the estimator of kth parameter α_k.

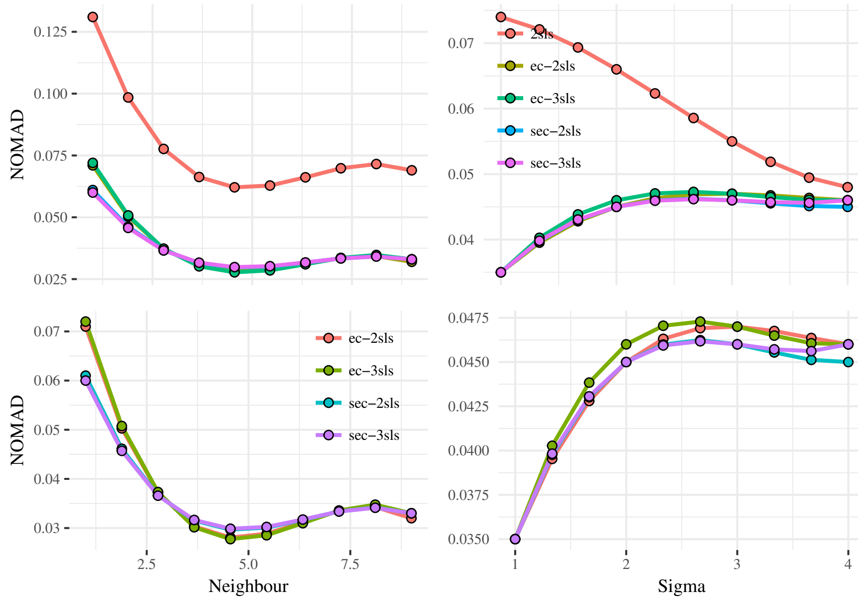

Figure 1: Estimators Average NOMAD evolution under covariance and spatial matrix.

As a supplement to $RMSE^*$ for each structural parameter, we employ two other comprehensive criteria proposed by Sasser (1969). The normalized mean absolute deviation (NOMAD) and normalized root mean square deviation (NORMSQD). These measures were also used by Baltagi (1984) in his Monte Carlo experiments. Specifically, NOMAD is defined as

$$NOMAD(\hat{\alpha}) = \frac{1}{RK} \sum_{k=1}^{K} \sum_{r=1}^{R} \left| \frac{\hat{\alpha}_{k,r} - \alpha_k}{\alpha_k} \right|$$

where K is the number of parameters, i.e., the dimension of parameter vector, R is the number of replications, $\hat{\alpha}_{k,r}$ is the estimator of kth parameter k in rth replication. Since $NORMSQD$ relies on moments as well, we will also use quantiles instead to adjust for this criterion. Therefore, $NORMSQD$ becomes

$$NORMSQD^*(\hat{\alpha}) = \left[\frac{1}{K} \sum_{k=1}^{K} \frac{\left[bias^2(\hat{\alpha}_k) + \left(\frac{IQ(\hat{\alpha}_k)}{1.35} \right)^2 \right]}{\hat{\alpha}_k^2} \right]^{1/2}$$

where $bias$ and IQ are defined similarly to those in $RMSE$. For simplicity of notation, we still use RMSE and NORMSQD in the text when using these adjusted measures.

6.3 Results

6.3.1 Changes in the Variance-Covariance matrix

Tables 4, 5, 6 and 7 show the bias, the standard deviation, the RMSE, the NOMAD and the NORMSQD based on 1000 replications. The structural parameters $(\alpha_1, \beta_{11}, \beta_{12}, \alpha_2, \beta_{21}, \beta_{22})$ take the values $(-0.5, -2, 1.5, -4, -3, 1.8)$, the spatial coefficients (ρ_1, ρ_2) are fixed at $(-0.8, -0.3)$, and the weighting matrix is fixed at W_3. The four tables differ only in the degree of heterogeneity in the individual effects and in the cross-equation correlation[5].
For the first three (03) usual estimators, with not surprise, OLS gives the largest NOMAD and NORMSQD. This is due to it inconsistency in a panel data (see Kelejian et al., 2004; Baltagi and Deng, 2015, for similar results). In contrast, when endogeneity is taken care of, i.e., we applied 2SLS, NOMAD gives an average gain of around 68.73% over OLS. In addition of 2SLS, when we swipe off all the specific effects, i.e., FE-2SLS is applied, now NOMAD exhibits an average gain of around 40.74% over 2SLS. As we move from V1 to V4, i.e. the variances of the individual effects decrease, 2SLS shows smaller RMSE as well as NOMAD and NORMSQD than FE-2SLS. For example, in Table 7 the NORMSQD gives an average gain of 33.01% over FE-2SLS (one can revisit Baltagi and Deng, 2015, for similar results).
Next, we compare 2SLS, FE-2SLS, EC-2SLS and EC-3SLS. As we see, when the variance covariance of the idiosyncratic term is small (Table 4), according to NOMAD and NORMSQD, EC-(3SLS and 2SLS) give better results than FE and 2SLS. According to standard deviation and RMSE, EC-3SLS exhibits better results than EC-2SLS. This is not surprising considering the fact that 2SLS does not take into account simultaneity. EC-2SLS NOMAD (0.35) is less than those of FE-2SLS (0.36). As the variance of individual effects decrease, the inequality between FE-2SLS and EC (2SLS and 3SLS) increases. For example in Table 7, considering EC-2SLS, NOMAD exhibits an average gain of 36.11% than FE-2SLS.
Now we compare classical error component estimators and spatial (AR) error component estimators. When the covariance matrix of individual effects are high (Table 4), results reveal that even if NOMADs are equal, NORMSQD, standard deviations and RMSE of spatial error components are less than those of classical error component estimators. For example, EC-2SLS normalized root mean square deviation is 0.052 while its equivalent in the spatial approach (SEC-2SLS) is 0.051. As the covariance matrix of individual effect decreases, this inequality holds. When we move from SEC-2SLS to SEC-3SLS, three stage least squares give better results than spatial two stage least squares. Indeed, even if RMSE and NORMSQD are close, we have a tiny gain in efficiency in favor of SEC-3SLS of the RMSE (respectively standard deviation) of each structural parameters. See Figure 1 column *Sigma* for a graphical appreciation.
Finally, we compare spatial (MA) error component and spatial (AR) error component estimators. As we see in Table 4 even if standard deviations and RMSE are closes, NOMAD exhibits an average gain of around 8.57% while NORMSQD gives 7.84% over spatial autoregressive error components. This difference in favor of SMA can be explained by shock affectations. Indeed, in SMA, spatial shock propagations are local's while in SAR process

[5] This respectively correspond to the following covariance setup: V1, V2, V3 and V4.

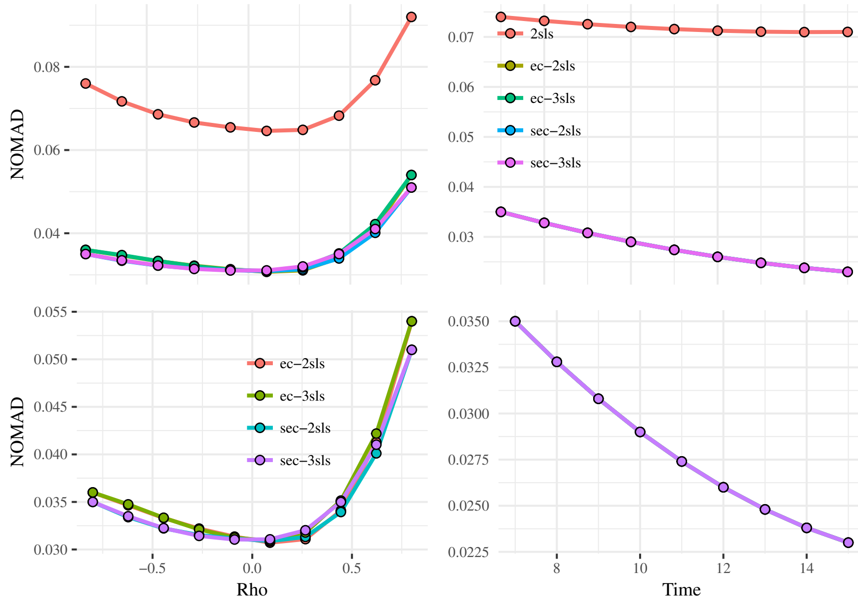

Figure 2: Estimators Average NOMAD evolution under time and spatial coefficient

they are global's.

6.3.2 Change in the number of Neighbours

Tables 8, 9 and 10 differ from Table 4 in the number of neighbours J. In Table 8 $J = 1$, in Table 4 $J = 3$, in Table 9 $J = 7$ and $J = 9$ in Table 10. The structural parameters $(\alpha_1, \beta_{11}, \beta_{12}, \alpha_2, \beta_{21}, \beta_{22})$ take the value $(-0.5, -2, 1.5, -4, -3, 1.8)$, the spatial coefficients (ρ_1, ρ_2) are fixed at $(-0.8, -0.3)$, and the variance covariance matrix design is fixed at V_1. This means that, the four tables differ only in the degree of sparseness. The non-zero rate of spatial matrices W_1, W_3, W_7 and W_9 are respectively 7.68%, 22.08%, 47.04% and 57.6% (see their representation in Figure 7).
For one neighbour ahead (W_1), results reveal that spatial three stage least squares give better results than spatial two stages. The average gain in effeciency of NOMAD is 1.64%. And as the number of neighbours increase i.e. as we move from W_1 to W_9, this inequality holds (see Figure 1 column Neighbour).
Comparing SAR and SMA approaches, as we move from Table 8 to Table 10, their NOMAD and NORMSQD become more close. It seems that when the number of neighbours increases, this tend to vanish to effect between this two approaches.

6.3.3 Change in time

Tables 4, 11 and 12 deal with the change in temporal size. Indeed, the first table consider $T = 7$ the second $T = 10$ and the third $T = 15$. The structural parameters $(\alpha_1, \beta_{11}, \beta_{12}, \alpha_2, \beta_{21}, \beta_{22})$, the coefficient of spatial dependence of each equation (ρ_1, ρ_2) take

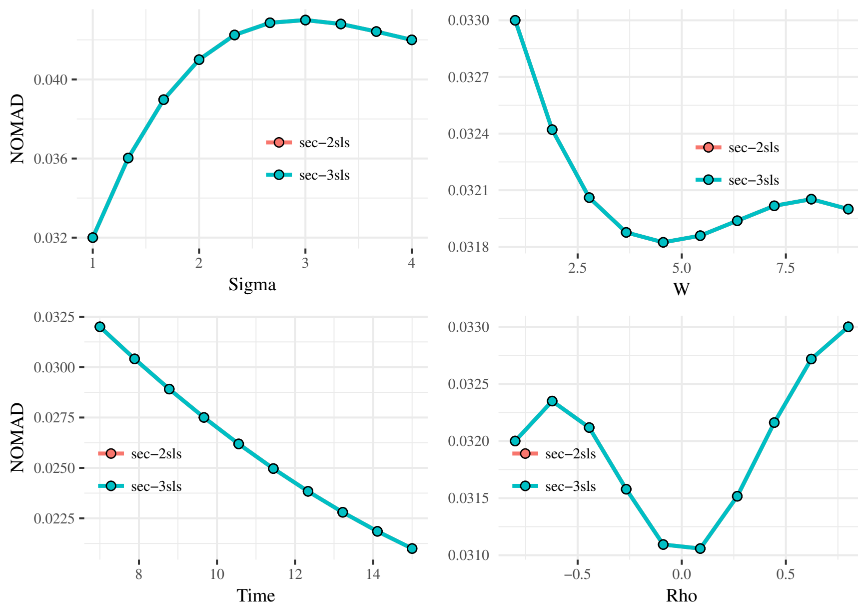

Figure 3: Estimators Average NOMAD evolution in a SMA case

the same values as above and the variance covariance matrix design is fixed at V_1. Comparing the results from Table 4 to 11 we see that an increase of temporal size leads to (i) a decrease in NOMAD and NORMSQD in all estimators (see Figure 2 column $Time$); (ii) SEC-3SLS are significantly better than SEC-2SLS only in standard deviation and RMSE of structural parameters.

6.3.4 Change in coefficient of spatial dependence

Tables 13 to 14 show the bias, the standard deviation, the RMSE, the NOMAD and the NORMSQD based on 1000 replications[6]. The structural parameters $(\alpha_1, \beta_{11}, \beta_{12}, \alpha_2, \beta_{21}, \beta_{22})$ take the values $(-0.5, -2, 1.5, -4, -3, 1.8)$, the weighting matrix is fixed at W_3, we use the first covariance design V_1 and time dimension is fixed at 7. In this investigation, ρ_1 and ρ_2 take the same values. In the first table $\rho_l = -0.8$, in the second $\rho_l = -0.4$, in the third ρ_l is null, in the fourth $\rho_l = 0.4$ and in the fifth $\rho_l = -0.8$. As we see in Figure 2 column $\boldsymbol{\rho}$ when ρ_l is negative, RMSE, NOMAD and NORMSQD are progressively decreasing while when this coefficient changes sign, NOMAD and NORMSQD progressively increase. For $\rho_l = 0$, SEC-3SLS (respectively SEC-2SLS) and EC-3SLS (respectively EC-2SLS) gives the same standard deviation, RMSE, NOMAD and NORMSQD.

6.3.5 Estimation of the coefficient of spatial dependence

Table 2 contains results on a measure of dispersion relating to the small sample distribution of our GMM estimator of ρ_l for each sixteen (16) cases. We applied GMM procedure in

[6]We only report the first and the last tables. The others are available on request.

two approaches; in the first, we used residuals of EC-2SLS whereas in the second we used residuals for EC-3SLS. We note some points.
The absolute average biases are generally similar for $\hat{\rho}_{2SLS}$ and $\hat{\rho}_{3SLS}$. Error components 3SLS gives better results than 2SLS. Which means that taking into account of simultaneity can slightly improve the quality of our GMM estimator. For all the weighting matrices considered in our experiment, as the number of neighbours increases, RMSE proportionaly increases. When we move from $\rho_l = -0.8$ to $\rho_l = 0.8$ each RMSE of W_J progressively decreases to its smallest value. We also study sample properties of our GMM estimators $\hat{\rho}_{2SLS}$ and $\hat{\rho}_{3SLS}$ for $J = 3$ under variance covariance designs. Results are similar from the spatial case (see Figure 4 for a better visualization).

Table 2: Bias and RMSEs of the estimator ρ_l

Parameter values		**EC-2SLS**		**EC-3SLS**	
J	ρ	Bias	ρ_{2SLS}	Bias	ρ_{3SLS}
One Neighbour					
1	-0.8	0.043	0.063	0.043	0.063
1	-0.4	0.013	0.072	0.012	0.071
1	0.4	0.008	0.069	0.008	0.068
1	0.8	0.037	0.059	0.036	0.058
Three Neighbours					
3	-0.8	0.023	0.182	0.021	0.183
3	-0.4	0.001	0.170	0.000	0.173
3	0.4	0.016	0.108	0.016	0.108
3	0.8	0.017	0.054	0.016	0.053
Seven Neighbours					
7	-0.8	0.029	0.320	0.031	0.313
7	-0.4	0.044	0.271	0.045	0.272
7	0.4	0.040	0.152	0.039	0.149
7	0.8	0.022	0.063	0.022	0.064
Nine Neighbours					
9	-0.8	0.053	0.365	0.055	0.356
9	-0.4	0.064	0.319	0.065	0.306
9	0.4	0.050	0.165	0.050	0.163
9	0.8	0.025	0.068	0.025	0.068

7 Health Expenditure and Real Income in SSA

7.1 The Model

No formal theory is available that predicts per capita health care expenditure (HCE) and economic growth. However, Parkin et al. (1987) developed a theory of national health expenditure in order to model the purchasing behavior at individual as well as at family levels The literature on economic growth suggests that investment in human capital particularly

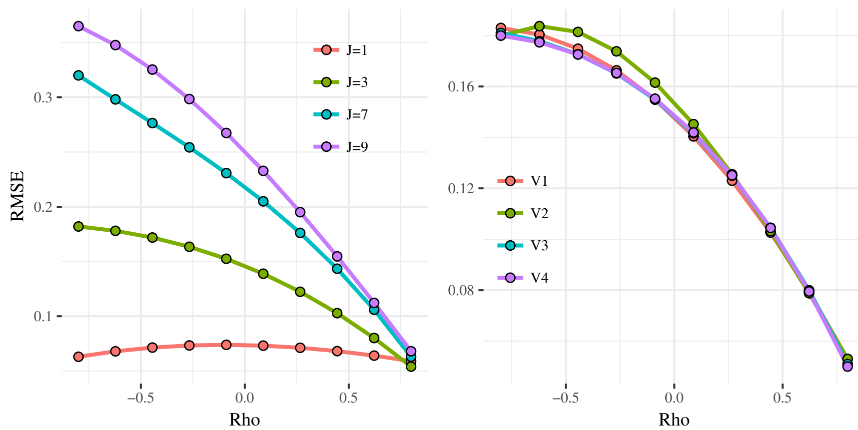

Figure 4: Spatial coefficient RMSE evolution

on health promotes the economic growth. Barro and Sala-i Martin (1995) argued that factors such as economic conditions, the level of economic development, and human and physical capital stock are the driving force of economic growth. Solow (1956) suggests that simultaneous increase in the level of physical and human capital causes per capita GDP and increased expenditure on health spurs the economic growth. To examine the dynamics of health expenditures, non-economic factors such as elderly population age of 65 years and above have play an important role. Hitiris and Posnett (1992) and O'Connell (1996) found that a percent of elderly population age of 65 years and above had a significant correlation with per capita health care expenditure. They used elderly age of 65 and above to show high potential health care. It can be argued that there exists bidirectional causality relationship between HCE and economic growth. On one hand, economic growth exerts positive impacted on HCE and on the other hand, HCE causes economic growth through its impact on the labor productivity. Taking the lead from the above analytical framework, we specify the following spatial simultaneous empirical model:

$$\begin{cases} y_{ti} = \beta_{01} + \alpha_1 h_{ti} + \beta_1 f_{ti} + \beta_2 k_{ti} + \beta_3 op_{ti} + \upsilon_{1,ti} \\ h_{ti} = \beta_{02} + \alpha_2 y_{ti} + \beta_4 pub_{ti} + \beta_5 old_{ti} + \beta_6 young_{ti} + \upsilon_{2,ti} \end{cases} \tag{49}$$

where h_{ti} and y_{ti}, the dependent variables of the system, respectively indicate, per capita health care expenditure and real GDP per capita for the ith country at time t; the exogenous variables of the model f, k, *pub*, *old* and *young* respectively indicate labor force, physical capital, trade openness, public expenditure on health care, the dependency rates for old and young people, defined as the population aged 65 and over divided by the population aged 15–64, and the population aged 0–14 divided by the population aged 15–64. All variables in equation (49) are expressed in natural logarithm. The structural disturbances for each equation follows a SAR process defined as in equation (2) with $l = \{1, 2\}$.

$$\upsilon_l = \Lambda_l \epsilon_l = (I - \rho_l W)^{-1} \epsilon_l = A_l^{-1} \epsilon_l \tag{50}$$

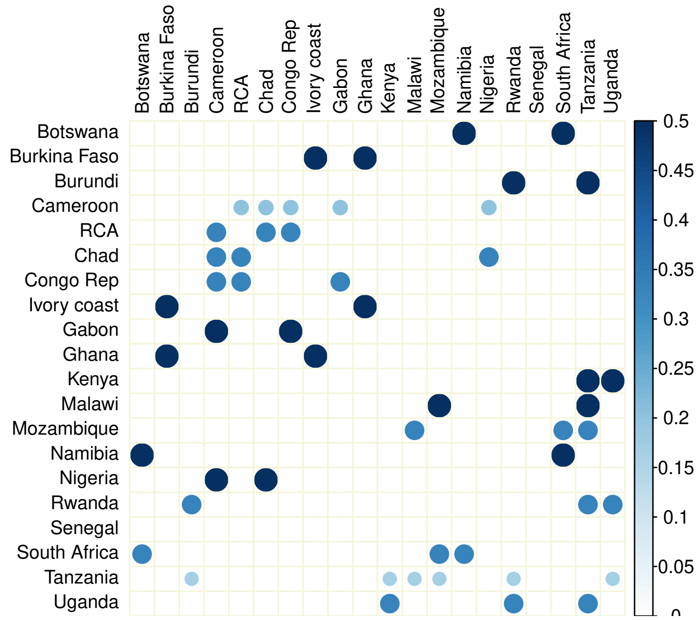

Figure 5: Spatial matrix of 20 SSA countries

Where the nonnegative matrix $W_N = (\omega_{ij})$, known as spatial weights matrix, provides information on the neighborhood linkages among Sub-Saharan African countries. In this study, we define neighborliness via a contiguity criterion, and assign $\omega_{ij} = 1$ when country i and j share a common border or vertex, and $\omega_{ij} = 0$ otherwise. This spatial matrix W_N gives a non-zero rate of 13% (see figure 5). The innovations ϵ_l follows a one-way error component model defined as in equation (3)

$$\epsilon_l = Z_\eta \eta_l + \xi_l$$

7.2 Data

We used annual data of 20 SSA countries over the period 1995 to 2015. The data comes from the World Development Indicators as published by the World Bank (2017). The selected countries under study and time span are dictated by data availability. Figure 6 displays the evolution of per capita HCE and per capita GDP for Sub-Saharan African countries.
We first made a preliminary exploratory data analysis this means: check whether our variables are nonstationary, then test their cointegrating properties and therefore, if they are linked in the long-run[7].

[7] We found that our variables are nonstationary in level, i.e, they are I(1). And the Johansen cointegration test reveals that they are cointegrated. For space requirement, results are not plot here but are availabe in request.

7.3 Empirical Comments

Table 3 shows results from: Ordinary Least Squares (column 1), Two Stage Least Squares (column 2), Least Square Dummy Variable 2SLS (column 3), Error component 2SLS (column 4), Error Components 3SLS (column 5), Spatial Error Component 2SLS (column 6) and Spatial Error Component Three Stage Least Squares (column 7) estimations when income is the dependent variable in the regression (Equation 1), as well as when health expenditure is the dependent variable (Equation 2).
Equation 1 reveals the following results. The impact of Health Care Expenditure on per capita income is negative in SSA; so an increment of 1% in HCE leads to a reduction of 0.037% in per capita GDP. Capital Stock and Trade Openness positively and significantly affect per capita Income; for example when we focus on 3EC-3SLS, an increase of 1% on capital stock (respectively Trade Openness) leads *ceteris paribus* in long run, to an increment of 0.305% (respectively 0.18%) of per capita GDP. Interestingly, per capita GDP is negatively affected by Labor force. This can be explaining by the properties of real Income. Indeed, our dependent variable is per capita GDP this means gross domestic product of a country divide by the population of the same year. So a negative impact of Labor force on per capita GDP in Sub-Saharan Africa means that the total quantity value-added by Labor force is less than the quantity absorbed by the excess of the population in the same period. In others words, even if Labor force play a positive role on GDP, if the natural growth is significantly high, it can imply a reduction of per capita GDP.

In the second equation 2 health expenditure is the dependent variable. As we can see, health care expenditure is positively impact by per capita GDP and the magnitude vary according to the estimator used. Indeed, a 1% increase in per capita real income leads in long run to an increase of 0.894% in health care expenditure for 3EC-3SLS and 1.431% in HCE for 2EC-3SLS. Public expenditure positively and significantly affects health expenditure while old ratio dependency rate and young ratio dependency rate negatively affect Health care expenditure.

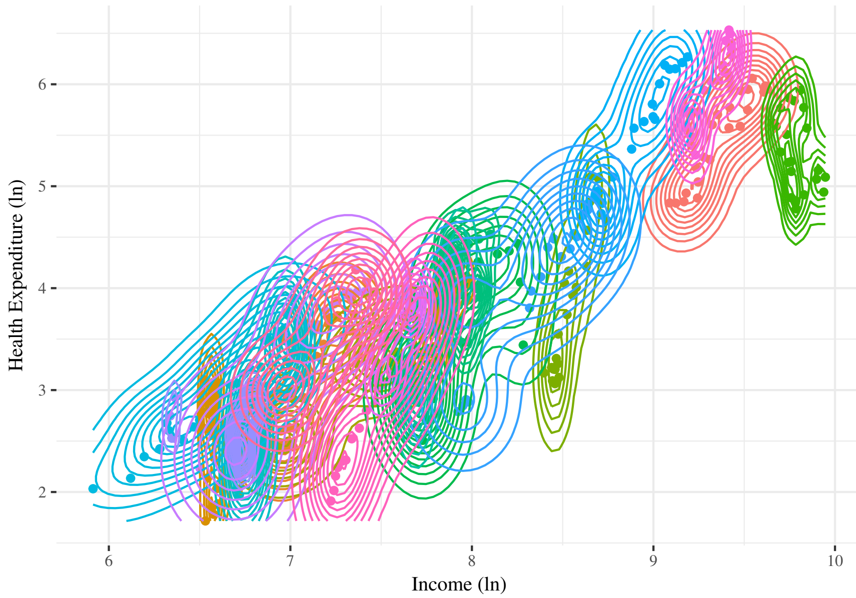

Figure 6: Health Expenditure and Income versus in SSA

Table 3: Health and Income in SSA SPE-SAR Estimates [1]

Variables	OLS	2SLS	LSDV	EC2SLS	EC3SLS	SEC2SLS	SEC3SLS
Real Income							
Constant	8.062*	10.664*	0.349	8.445*	10.459*	8.734*	7.048*
	(0.433)	(0.702)	(190.776)	(3.279)	(2.99)	(3.242)	(3.258)
Health	0.324*	-0.101	-0.099	-0.024	-0.003	-0.019	-0.037
	(0.032)	(0.087)	(0.056)	(0.069)	(0.064)	(0.068)	(0.068)
Labour	-0.463*	-0.804*	0.336*	-0.281	-0.45*	-0.3	-0.196
	(0.034)	(0.075)	(0.113)	(0.234)	(0.213)	(0.231)	(0.232)
Capital	0.424*	0.789*	0.201*	0.288*	0.337*	0.289*	0.305*
	(0.033)	(0.077)	(0.024)	(0.031)	(0.029)	(0.03)	(0.03)
Openness	0.255*	0.325*	0.112*	0.181*	0.203*	0.18*	0.18*
	(0.049)	(0.06)	(0.034)	(0.034)	(0.032)	(0.034)	(0.032)
Health Expenditure							
Constant	-5.653*	-6.239*	-16.314	-11.155*	-9.351*	-12.004*	-6.143*
	(0.472)	(0.516)	(14.198)	(0.917)	(0.887)	(0.868)	(1.01)
Income	1.005*	1.064*	2.264*	1.784*	1.663*	1.886*	1.327*
	(0.032)	(0.038)	(0.094)	(0.078)	(0.073)	(0.077)	(0.091)
Public	0.583*	0.607*	0.53*	0.673*	0.682*	0.646*	0.528*
	(0.044)	(0.045)	(0.044)	(0.042)	(0.042)	(0.039)	(0.041)
Old	-0.353*	-0.411*	-0.604*	-0.157	0.134	-0.171	0.256
	(0.125)	(0.127)	(0.21)	(0.233)	(0.224)	(0.213)	(0.205)
Young	-0.486*	-0.242	-1.28*	-0.227	-0.226	-0.432	-0.043
	(0.184)	(0.204)	(0.264)	(0.262)	(0.267)	(0.242)	(0.255)

[1] Simultaneous Panel Equation with Spatial Autoregressive Error.
[2] The number in parentheses denotes the standard deviation.
[3] * Denote significance at 5% of the parameter.

8 Conclusion

This paper develops estimation for a simultaneous panel data with spatially autocorrelated error componenent. For the disturbance, we considered SAR process developped by Kelejian and Prucha (2004) in which the global effect shock occurs because it is transmitted also to location that are "neigbours of neighbours" via the power of the spatial matrix. We also consider SMA process developped by Fingleton (2008) in which a shock at a specific location will only affect the directly interacting location. We derive a limited information estimator, termed SEC2SLS estimator, and a full information estimator, termed SEC3SLS. To derive our spatial error component estimators, we need a procedure which both resolve the endogeneity problem and spatial correlation. Thus we combine instrumental variable and spatial GLS estimator. And the coefficient of the spatial dependence of each equation is therefore derived by using GMM procedure.

The purpose of our Monte Carlo experiment were threefolds: Firstly, we study the small sample behavior of our estimators that can handle endogeneity, spatial error correlation and random individual effects in function of spatial coefficient, contiguity matrix, variance covariance of specific component and the increase of time. These estimators are compared with those that may ignore one or more of these symptoms. Secondly, we also investigate the gains in efficiency by comparing SEC to others estimators. Thirdly, we study the sample properties of the spatial component ρ_l in limited and full information cases. Results suggesting many conclusions. Our estimators are consistent. According to RMSE and standard deviation, SEC3SLS is better than SEC2SLS. When we estimate the coefficient of spatial dependence it seems better to use IV estimator that takes into account simultaneity. This means that when it is possible use EC3SLS in lieu of EC2SLS.

Finally, we apply these estimators to real data of 20 Sub-Saharan African countries . We used these set of estimators to evaluate the modification of the magnitude in the model of health care expenditure and per capita real income.

In future research it should be of interest to extend the analysis of this paper to the case that contains spatial lag and spatially autocorrelated error components.

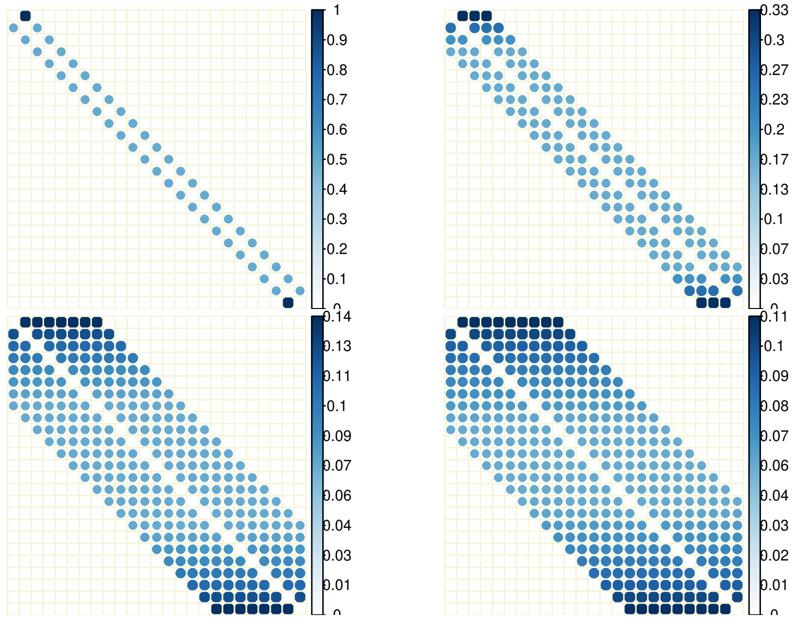

Figure 7: Spatial Matrices under different Neighbourhood

A Appendix

Table 4: Efficient Criteria under Covariance V1

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.106	0.851	0.64	0.003	0.002	0	0.194
	(0.01)	(0.095)	(0.077)	(0.021)	(0.061)	(0.056)	(0.46)
	[0.11]	[0.877]	[0.662]	[0.046]	[0.13]	[0.119]	
2SLS	0.001	0.005	0	0.001	0.004	0.003	0.074
	(0.018)	(0.153)	(0.121)	(0.022)	(0.062)	(0.057)	(0.108)
	[0.038]	[0.332]	[0.25]	[0.048]	[0.129]	[0.118]	
LSDV	0.001	0.01	0.003	0	0.002	0.002	0.036
	(0.019)	(0.163)	(0.129)	(0.023)	(0.067)	(0.061)	(0.052)
	[0.018]	[0.153]	[0.126]	[0.021]	[0.064]	[0.062]	
One Way Error Components							
2SLS	0.006	0.048	0.04	0	0.002	0.002	0.035
	(0.017)	(0.144)	(0.115)	(0.021)	(0.061)	(0.056)	(0.052)
	[0.018]	[0.153]	[0.124]	[0.021]	[0.064]	[0.063]	
3SLS	0.007	0.052	0.043	0.001	0.009	0.005	0.035
	(0.016)	(0.142)	(0.111)	(0.021)	(0.059)	(0.051)	(0.053)
	[0.018]	[0.158]	[0.123]	[0.02]	[0.062]	[0.058]	
SAR Error Components							
S2SLS	0.006	0.046	0.039	0	0.001	0.002	0.035
	(0.017)	(0.145)	(0.115)	(0.021)	(0.061)	(0.056)	(0.051)
	[0.018]	[0.152]	[0.12]	[0.021]	[0.064]	[0.062]	
S3SLS	0.006	0.05	0.041	0.001	0.009	0.005	0.035
	(0.016)	(0.142)	(0.112)	(0.021)	(0.059)	(0.051)	(0.052)
	[0.018]	[0.154]	[0.122]	[0.02]	[0.061]	[0.058]	
SMA Error Components							
S2SLS	0.005	0.037	0.032	0	0.001	0.002	0.032
	(0.016)	(0.138)	(0.11)	(0.021)	(0.06)	(0.055)	(0.047)
	[0.016]	[0.137]	[0.112]	[0.02]	[0.063]	[0.062]	
S3SLS	0.005	0.04	0.034	0.001	0.007	0.004	0.032
	(0.016)	(0.136)	(0.107)	(0.021)	(0.058)	(0.05)	(0.047)
	[0.017]	[0.141]	[0.108]	[0.02]	[0.06]	[0.056]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 5: Efficient Criteria under Covariance V2

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.108	0.864	0.649	0.003	0.002	0	0.195
	(0.01)	(0.095)	(0.078)	(0.021)	(0.062)	(0.057)	(0.465)
	[0.11]	[0.885]	[0.663]	[0.041]	[0.112]	[0.105]	
2SLS	0.001	0.005	0	0.001	0.004	0.003	0.066
	(0.018)	(0.154)	(0.122)	(0.022)	(0.063)	(0.057)	(0.096)
	[0.033]	[0.295]	[0.224]	[0.044]	[0.116]	[0.107]	
LSDV	0.002	0.015	0.005	0	0.003	0.003	0.051
	(0.027)	(0.232)	(0.184)	(0.033)	(0.095)	(0.086)	(0.073)
	[0.026]	[0.221]	[0.176]	[0.03]	[0.091]	[0.088]	
One Way Error Components							
2SLS	0.009	0.072	0.059	0	0.001	0.002	0.045
	(0.021)	(0.182)	(0.144)	(0.027)	(0.077)	(0.07)	(0.069)
	[0.023]	[0.201]	[0.16]	[0.026]	[0.081]	[0.078]	
3SLS	0.01	0.076	0.062	0.001	0.013	0.008	0.046
	(0.02)	(0.177)	(0.139)	(0.026)	(0.075)	(0.064)	(0.07)
	[0.024]	[0.207]	[0.162]	[0.026]	[0.08]	[0.073]	
SAR Error Components							
S2SLS	0.009	0.071	0.057	0	0.001	0.002	0.045
	(0.021)	(0.182)	(0.145)	(0.027)	(0.077)	(0.07)	(0.067)
	[0.023]	[0.193]	[0.156]	[0.026]	[0.08]	[0.077]	
S3SLS	0.009	0.074	0.059	0.001	0.013	0.007	0.045
	(0.02)	(0.178)	(0.14)	(0.026)	(0.074)	(0.064)	(0.068)
	[0.023]	[0.197]	[0.16]	[0.026]	[0.078]	[0.073]	
SMA Error Components							
S2SLS	0.007	0.059	0.049	0	0.001	0.001	0.041
	(0.02)	(0.174)	(0.138)	(0.026)	(0.076)	(0.069)	(0.06)
	[0.021]	[0.175]	[0.14]	[0.026]	[0.079]	[0.076]	
S3SLS	0.008	0.061	0.05	0.001	0.011	0.006	0.041
	(0.02)	(0.17)	(0.134)	(0.026)	(0.074)	(0.064)	(0.062)
	[0.022]	[0.18]	[0.147]	[0.026]	[0.077]	[0.072]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 6: Efficient Criteria under Covariance V3

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.11 (0.01) [0.112]	0.878 (0.095) [0.895]	0.66 (0.078) [0.671]	0.01 (0.019) [0.03]	0.014 (0.056) [0.081]	0.007 (0.052) [0.075]	0.193 (0.476)
2SLS	0.001 (0.018) [0.03]	0.004 (0.154) [0.262]	0 (0.122) [0.196]	0.001 (0.02) [0.028]	0.002 (0.057) [0.081]	0.003 (0.052) [0.075]	0.055 (0.083)
LSDV	0.002 (0.033) [0.031]	0.019 (0.287) [0.269]	0.006 (0.227) [0.212]	0 (0.04) [0.036]	0.003 (0.116) [0.111]	0.004 (0.106) [0.108]	0.062 (0.089)
One Way Error Components							
2SLS	0.01 (0.022) [0.025]	0.08 (0.193) [0.217]	0.064 (0.154) [0.174]	0 (0.025) [0.025]	0.001 (0.072) [0.076]	0 (0.065) [0.069]	0.047 (0.074)
3SLS	0.01 (0.022) [0.026]	0.081 (0.188) [0.221]	0.065 (0.147) [0.174]	0.001 (0.024) [0.025]	0.015 (0.068) [0.076]	0.01 (0.057) [0.067]	0.047 (0.074)
SAR Error Components							
S2SLS	0.01 (0.022) [0.025]	0.078 (0.193) [0.209]	0.062 (0.154) [0.171]	0 (0.025) [0.025]	0.001 (0.071) [0.078]	0 (0.065) [0.069]	0.046 (0.072)
S3SLS	0.01 (0.022) [0.025]	0.079 (0.188) [0.214]	0.063 (0.147) [0.169]	0.001 (0.024) [0.025]	0.014 (0.067) [0.075]	0.009 (0.057) [0.065]	0.046 (0.072)
SMA Error Components							
S2SLS	0.008 (0.021) [0.022]	0.067 (0.185) [0.194]	0.054 (0.148) [0.156]	0 (0.024) [0.025]	0.001 (0.07) [0.076]	0 (0.064) [0.068]	0.043 (0.066)
S3SLS	0.008 (0.021) [0.023]	0.067 (0.181) [0.192]	0.054 (0.142) [0.156]	0.001 (0.024) [0.025]	0.012 (0.067) [0.073]	0.008 (0.057) [0.064]	0.043 (0.066)

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 7: Efficient Criteria under Covariance V4

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.11	0.883	0.663	0.003	0.003	0	0.195
	(0.01)	(0.095)	(0.078)	(0.021)	(0.063)	(0.058)	(0.474)
	[0.112]	[0.892]	[0.669]	[0.029]	[0.084]	[0.076]	
2SLS	0	0.003	0	0.001	0.003	0.003	0.048
	(0.018)	(0.155)	(0.123)	(0.022)	(0.064)	(0.058)	(0.069)
	[0.024]	[0.212]	[0.162]	[0.031]	[0.085]	[0.077]	
LSDV	0.002	0.022	0.007	0	0.004	0.005	0.072
	(0.038)	(0.335)	(0.264)	(0.046)	(0.134)	(0.122)	(0.103)
	[0.036]	[0.31]	[0.245]	[0.042]	[0.128]	[0.125]	
One Way Error Components							
2SLS	0.009	0.071	0.056	0	0.001	0.001	0.046
	(0.021)	(0.183)	(0.146)	(0.027)	(0.078)	(0.071)	(0.072)
	[0.024]	[0.213]	[0.168]	[0.028]	[0.082]	[0.075]	
3SLS	0.009	0.074	0.059	0.001	0.015	0.009	0.046
	(0.021)	(0.179)	(0.14)	(0.027)	(0.075)	(0.065)	(0.072)
	[0.025]	[0.211]	[0.17]	[0.029]	[0.084]	[0.073]	
SAR Error Components							
S2SLS	0.009	0.069	0.055	0	0.001	0.001	0.045
	(0.021)	(0.182)	(0.145)	(0.027)	(0.078)	(0.071)	(0.071)
	[0.024]	[0.207]	[0.167]	[0.028]	[0.082]	[0.074]	
S3SLS	0.009	0.072	0.057	0.002	0.014	0.009	0.046
	(0.021)	(0.178)	(0.14)	(0.027)	(0.075)	(0.065)	(0.071)
	[0.024]	[0.211]	[0.165]	[0.029]	[0.083]	[0.071]	
SMA Error Components							
S2SLS	0.008	0.06	0.048	0	0.001	0.001	0.042
	(0.02)	(0.176)	(0.14)	(0.027)	(0.077)	(0.07)	(0.065)
	[0.022]	[0.188]	[0.154]	[0.027]	[0.081]	[0.073]	
S3SLS	0.008	0.062	0.05	0.002	0.013	0.008	0.042
	(0.02)	(0.172)	(0.136)	(0.026)	(0.074)	(0.064)	(0.065)
	[0.023]	[0.192]	[0.15]	[0.028]	[0.081]	[0.07]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 8: Efficient Criteria under Neighbour J=1

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.189	1.511	1.137	0.027	0.037	0.022	0.333
	(0.009)	(0.093)	(0.08)	(0.02)	(0.063)	(0.058)	(1.871)
	[0.192]	[1.542]	[1.158]	[0.053]	[0.138]	[0.131]	
2SLS	0.001	0.009	0.02	0.001	0.005	0.004	0.131
	(0.04)	(0.342)	(0.268)	(0.023)	(0.066)	(0.06)	(0.204)
	[0.069]	[0.578]	[0.481]	[0.051]	[0.137]	[0.13]	
LSDV	0.001	0.014	0.004	0	0.002	0.003	0.062
	(0.038)	(0.326)	(0.258)	(0.024)	(0.071)	(0.065)	(0.099)
	[0.034]	[0.309]	[0.247]	[0.023]	[0.067]	[0.065]	
One Way Error Components							
2SLS	0.026	0.201	0.161	0.001	0.004	0.003	0.071
	(0.03)	(0.267)	(0.213)	(0.022)	(0.064)	(0.059)	(0.126)
	[0.041]	[0.35]	[0.288]	[0.021]	[0.067]	[0.064]	
3SLS	0.026	0.204	0.162	0	0.016	0.009	0.072
	(0.03)	(0.259)	(0.204)	(0.022)	(0.062)	(0.054)	(0.127)
	[0.042]	[0.355]	[0.288]	[0.021]	[0.07]	[0.064]	
SAR Error Components							
S2SLS	0.019	0.148	0.119	0.001	0.003	0.003	0.061
	(0.032)	(0.283)	(0.226)	(0.022)	(0.064)	(0.059)	(0.105)
	[0.036]	[0.305]	[0.238]	[0.022]	[0.068]	[0.064]	
S3SLS	0.018	0.142	0.113	0	0.01	0.005	0.06
	(0.032)	(0.277)	(0.219)	(0.022)	(0.062)	(0.055)	(0.105)
	[0.035]	[0.303]	[0.242]	[0.022]	[0.065]	[0.061]	
SMA Error Components							
S2SLS	0.005	0.038	0.032	0	0.001	0.002	0.033
	(0.018)	(0.156)	(0.124)	(0.021)	(0.061)	(0.056)	(0.051)
	[0.017]	[0.149]	[0.121]	[0.021]	[0.064]	[0.06]	
S3SLS	0.005	0.038	0.032	0.001	0.006	0.003	0.033
	(0.018)	(0.153)	(0.121)	(0.021)	(0.059)	(0.052)	(0.051)
	[0.017]	[0.148]	[0.123]	[0.02]	[0.063]	[0.057]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 9: Efficient Criteria under Neighbour J=7

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.096	0.768	0.577	0	0.002	0.002	0.176
	(0.01)	(0.093)	(0.075)	(0.021)	(0.061)	(0.055)	(0.386)
	[0.099]	[0.793]	[0.595]	[0.046]	[0.128]	[0.118]	
2SLS	0.001	0.004	0	0.001	0.004	0.003	0.069
	(0.016)	(0.14)	(0.111)	(0.021)	(0.061)	(0.056)	(0.102)
	[0.035]	[0.31]	[0.238]	[0.047]	[0.126]	[0.117]	
LSDV	0.001	0.01	0.003	0	0.002	0.002	0.033
	(0.017)	(0.15)	(0.119)	(0.023)	(0.066)	(0.06)	(0.047)
	[0.017]	[0.139]	[0.114]	[0.021]	[0.063]	[0.063]	
One Way Error Components							
2SLS	0.005	0.039	0.033	0	0.001	0.002	0.033
	(0.015)	(0.134)	(0.107)	(0.021)	(0.06)	(0.055)	(0.049)
	[0.017]	[0.141]	[0.115]	[0.02]	[0.063]	[0.061]	
3SLS	0.006	0.044	0.036	0.001	0.008	0.005	0.033
	(0.015)	(0.132)	(0.103)	(0.021)	(0.058)	(0.05)	(0.049)
	[0.017]	[0.148]	[0.115]	[0.02]	[0.062]	[0.056]	
SAR Error Components							
S2SLS	0.005	0.04	0.034	0	0.001	0.002	0.033
	(0.015)	(0.134)	(0.106)	(0.021)	(0.06)	(0.055)	(0.049)
	[0.017]	[0.142]	[0.117]	[0.02]	[0.063]	[0.061]	
S3SLS	0.006	0.044	0.037	0.001	0.008	0.005	0.033
	(0.015)	(0.132)	(0.103)	(0.021)	(0.058)	(0.05)	(0.049)
	[0.017]	[0.147]	[0.114]	[0.02]	[0.062]	[0.056]	
SMA Error Components							
S2SLS	0.005	0.038	0.032	0	0.001	0.002	0.032
	(0.015)	(0.132)	(0.105)	(0.021)	(0.06)	(0.055)	(0.047)
	[0.016]	[0.136]	[0.113]	[0.02]	[0.062]	[0.06]	
S3SLS	0.005	0.042	0.035	0.001	0.008	0.004	0.032
	(0.015)	(0.13)	(0.102)	(0.02)	(0.058)	(0.05)	(0.047)
	[0.016]	[0.137]	[0.111]	[0.02]	[0.061]	[0.056]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 10: Efficient Criteria under Neighbour J=9

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.095	0.757	0.569	0	0.002	0.002	0.174
	(0.01)	(0.093)	(0.075)	(0.021)	(0.06)	(0.055)	(0.381)
	[0.098]	[0.783]	[0.592]	[0.045]	[0.127]	[0.119]	
2SLS	0.001	0.005	0	0.001	0.004	0.003	0.069
	(0.016)	(0.139)	(0.11)	(0.021)	(0.061)	(0.056)	(0.102)
	[0.035]	[0.309]	[0.241]	[0.048]	[0.128]	[0.118]	
LSDV	0.001	0.01	0.003	0	0.002	0.002	0.033
	(0.017)	(0.149)	(0.118)	(0.023)	(0.066)	(0.06)	(0.047)
	[0.017]	[0.14]	[0.113]	[0.021]	[0.063]	[0.063]	
One Way Error Components							
2SLS	0.005	0.038	0.032	0	0.001	0.002	0.032
	(0.015)	(0.133)	(0.105)	(0.021)	(0.06)	(0.055)	(0.049)
	[0.017]	[0.143]	[0.116]	[0.02]	[0.064]	[0.06]	
3SLS	0.005	0.042	0.036	0.001	0.008	0.005	0.033
	(0.015)	(0.13)	(0.102)	(0.021)	(0.058)	(0.05)	(0.049)
	[0.017]	[0.144]	[0.115]	[0.02]	[0.062]	[0.056]	
SAR Error Components							
S2SLS	0.005	0.039	0.033	0	0.001	0.002	0.033
	(0.015)	(0.132)	(0.105)	(0.021)	(0.06)	(0.055)	(0.049)
	[0.017]	[0.14]	[0.117]	[0.02]	[0.063]	[0.06]	
S3SLS	0.005	0.043	0.036	0.001	0.008	0.005	0.033
	(0.015)	(0.13)	(0.102)	(0.021)	(0.058)	(0.05)	(0.049)
	[0.017]	[0.145]	[0.116]	[0.019]	[0.062]	[0.056]	
SMA Error Components							
S2SLS	0.005	0.038	0.032	0	0.001	0.002	0.032
	(0.015)	(0.131)	(0.104)	(0.021)	(0.06)	(0.055)	(0.047)
	[0.016]	[0.136]	[0.112]	[0.02]	[0.063]	[0.061]	
S3SLS	0.005	0.042	0.035	0.001	0.008	0.005	0.032
	(0.015)	(0.129)	(0.101)	(0.02)	(0.058)	(0.05)	(0.048)
	[0.017]	[0.139]	[0.114]	[0.019]	[0.061]	[0.056]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 11: Efficient Criteria for Time T=10

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.106	0.85	0.641	0.004	0.001	0.004	0.194
	(0.009)	(0.079)	(0.065)	(0.017)	(0.051)	(0.047)	(0.463)
	[0.109]	[0.882]	[0.667]	[0.047]	[0.14]	[0.121]	
2SLS	0.001	0.009	0.002	0	0.005	0	0.072
	(0.015)	(0.127)	(0.101)	(0.018)	(0.052)	(0.047)	(0.107)
	[0.036]	[0.32]	[0.259]	[0.046]	[0.144]	[0.123]	
LSDV	0	0.001	0.002	0	0	0.003	0.029
	(0.015)	(0.132)	(0.105)	(0.019)	(0.055)	(0.05)	(0.044)
	[0.015]	[0.134]	[0.106]	[0.019]	[0.058]	[0.049]	
One Way Error Components							
2SLS	0.004	0.035	0.027	0	0.001	0.003	0.029
	(0.014)	(0.122)	(0.097)	(0.018)	(0.052)	(0.047)	(0.045)
	[0.015]	[0.135]	[0.106]	[0.019]	[0.056]	[0.049]	
3SLS	0.005	0.039	0.029	0.001	0.008	0.003	0.029
	(0.014)	(0.12)	(0.094)	(0.018)	(0.05)	(0.043)	(0.045)
	[0.015]	[0.135]	[0.107]	[0.02]	[0.056]	[0.047]	
SAR Error Components							
S2SLS	0.004	0.034	0.025	0	0.001	0.003	0.029
	(0.014)	(0.122)	(0.097)	(0.018)	(0.052)	(0.047)	(0.043)
	[0.015]	[0.132]	[0.102]	[0.019]	[0.056]	[0.049]	
S3SLS	0.005	0.037	0.028	0.001	0.007	0.003	0.029
	(0.014)	(0.12)	(0.094)	(0.018)	(0.05)	(0.043)	(0.044)
	[0.015]	[0.135]	[0.105]	[0.02]	[0.057]	[0.047]	
SMA Error Components							
S2SLS	0.003	0.027	0.02	0	0.001	0.003	0.027
	(0.013)	(0.115)	(0.092)	(0.017)	(0.051)	(0.046)	(0.04)
	[0.014]	[0.121]	[0.095]	[0.019]	[0.055]	[0.048]	
S3SLS	0.004	0.03	0.021	0.001	0.006	0.002	0.027
	(0.013)	(0.114)	(0.09)	(0.017)	(0.049)	(0.042)	(0.04)
	[0.013]	[0.123]	[0.095]	[0.02]	[0.055]	[0.046]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 12: Efficient Criteria for Time, T=15

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.106	0.85	0.641	0.004	0.001	0.004	0.194
	(0.009)	(0.079)	(0.065)	(0.017)	(0.051)	(0.047)	(0.463)
	[0.109]	[0.882]	[0.667]	[0.047]	[0.14]	[0.121]	
2SLS	0.001	0.009	0.002	0	0.005	0	0.072
	(0.015)	(0.127)	(0.101)	(0.018)	(0.052)	(0.047)	(0.107)
	[0.036]	[0.32]	[0.259]	[0.046]	[0.144]	[0.123]	
LSDV	0	0.001	0.002	0	0	0.003	0.029
	(0.015)	(0.132)	(0.105)	(0.019)	(0.055)	(0.05)	(0.044)
	[0.015]	[0.134]	[0.106]	[0.019]	[0.058]	[0.049]	
One Way Error Components							
2SLS	0.004	0.035	0.027	0	0.001	0.003	0.029
	(0.014)	(0.122)	(0.097)	(0.018)	(0.052)	(0.047)	(0.045)
	[0.015]	[0.135]	[0.106]	[0.019]	[0.056]	[0.049]	
3SLS	0.005	0.039	0.029	0.001	0.008	0.003	0.029
	(0.014)	(0.12)	(0.094)	(0.018)	(0.05)	(0.043)	(0.045)
	[0.015]	[0.135]	[0.107]	[0.02]	[0.056]	[0.047]	
SAR Error Components							
S2SLS	0.004	0.034	0.025	0	0.001	0.003	0.029
	(0.014)	(0.122)	(0.097)	(0.018)	(0.052)	(0.047)	(0.043)
	[0.015]	[0.132]	[0.102]	[0.019]	[0.056]	[0.049]	
S3SLS	0.005	0.037	0.028	0.001	0.007	0.003	0.029
	(0.014)	(0.12)	(0.094)	(0.018)	(0.05)	(0.043)	(0.044)
	[0.015]	[0.135]	[0.105]	[0.02]	[0.057]	[0.047]	
SMA Error Components							
S2SLS	0.003	0.027	0.02	0	0.001	0.003	0.027
	(0.013)	(0.115)	(0.092)	(0.017)	(0.051)	(0.046)	(0.04)
	[0.014]	[0.121]	[0.095]	[0.019]	[0.055]	[0.048]	
S3SLS	0.004	0.03	0.021	0.001	0.006	0.002	0.027
	(0.013)	(0.114)	(0.09)	(0.017)	(0.049)	(0.042)	(0.04)
	[0.013]	[0.123]	[0.095]	[0.02]	[0.055]	[0.046]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 13: Efficient Criteria for ρ_l=-0.8

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.105	0.844	0.634	0.001	0.006	0.004	0.194
	(0.01)	(0.096)	(0.078)	(0.023)	(0.069)	(0.063)	(0.455)
	[0.109]	[0.873]	[0.658]	[0.051]	[0.143]	[0.135]	
2SLS	0.001	0.005	0	0.001	0.005	0.003	0.076
	(0.018)	(0.153)	(0.121)	(0.024)	(0.069)	(0.063)	(0.109)
	[0.038]	[0.333]	[0.249]	[0.053]	[0.149]	[0.136]	
LSDV	0.001	0.01	0.003	0	0.003	0.003	0.037
	(0.019)	(0.163)	(0.129)	(0.026)	(0.075)	(0.068)	(0.052)
	[0.018]	[0.154]	[0.126]	[0.023]	[0.073]	[0.07]	
One Way Error Components							
2SLS	0.006	0.047	0.039	0	0.002	0.002	0.036
	(0.017)	(0.144)	(0.115)	(0.023)	(0.068)	(0.062)	(0.053)
	[0.018]	[0.152]	[0.124]	[0.023]	[0.07]	[0.067]	
3SLS	0.007	0.052	0.043	0.001	0.01	0.006	0.036
	(0.016)	(0.142)	(0.111)	(0.023)	(0.066)	(0.056)	(0.053)
	[0.019]	[0.159]	[0.123]	[0.023]	[0.068]	[0.065]	
SAR Error Components							
S2SLS	0.006	0.046	0.038	0	0.002	0.002	0.035
	(0.017)	(0.145)	(0.115)	(0.023)	(0.068)	(0.062)	(0.052)
	[0.018]	[0.151]	[0.12]	[0.023]	[0.069]	[0.067]	
S3SLS	0.006	0.05	0.041	0.001	0.009	0.005	0.035
	(0.016)	(0.142)	(0.112)	(0.023)	(0.066)	(0.057)	(0.052)
	[0.018]	[0.154]	[0.121]	[0.023]	[0.068]	[0.063]	
SMA Error Components							
S2SLS	0.005	0.037	0.032	0	0.002	0.002	0.032
	(0.016)	(0.138)	(0.11)	(0.022)	(0.065)	(0.059)	(0.048)
	[0.017]	[0.137]	[0.112]	[0.021]	[0.065]	[0.062]	
S3SLS	0.005	0.04	0.034	0.001	0.007	0.004	0.032
	(0.016)	(0.135)	(0.106)	(0.022)	(0.063)	(0.054)	(0.047)
	[0.017]	[0.14]	[0.108]	[0.021]	[0.062]	[0.059]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

Table 14: Efficient Criteria for ρ_l=0.8

Method	α_1	β_{11}	β_{12}	α_2	β_{21}	β_{22}	Nomad
Value	-0.5	-2.0	1.5	-4	-3.0	1.8	Normsq
Usual Estimators							
OLS	0.137	1.099	0.824	0	0.002	0.002	0.252
	(0.011)	(0.101)	(0.084)	(0.029)	(0.088)	(0.081)	(0.715)
	[0.139]	[1.118]	[0.833]	[0.068]	[0.174]	[0.153]	
2SLS	0.001	0.005	0.006	0	0.003	0.002	0.092
	(0.023)	(0.201)	(0.16)	(0.031)	(0.09)	(0.082)	(0.127)
	[0.042]	[0.373]	[0.302]	[0.066]	[0.173]	[0.154]	
LSDV	0.002	0.015	0.006	0.002	0.002	0	0.057
	(0.029)	(0.252)	(0.199)	(0.039)	(0.115)	(0.105)	(0.082)
	[0.03]	[0.25]	[0.192]	[0.036]	[0.11]	[0.102]	
One Way Error Components							
2SLS	0.012	0.094	0.072	0.002	0.003	0.001	0.054
	(0.023)	(0.205)	(0.163)	(0.034)	(0.098)	(0.089)	(0.082)
	[0.028]	[0.236]	[0.189]	[0.033]	[0.101]	[0.096]	
3SLS	0.013	0.103	0.078	0.004	0.025	0.015	0.054
	(0.023)	(0.197)	(0.155)	(0.033)	(0.093)	(0.08)	(0.084)
	[0.029]	[0.245]	[0.189]	[0.033]	[0.104]	[0.084]	
SAR Error Components							
S2SLS	0.009	0.074	0.06	0.001	0.001	0.001	0.051
	(0.024)	(0.211)	(0.168)	(0.034)	(0.099)	(0.091)	(0.079)
	[0.029]	[0.231]	[0.183]	[0.032]	[0.095]	[0.088]	
S3SLS	0.01	0.08	0.062	0.003	0.019	0.012	0.051
	(0.024)	(0.206)	(0.162)	(0.034)	(0.096)	(0.083)	(0.08)
	[0.028]	[0.229]	[0.187]	[0.032]	[0.099]	[0.081]	
SMA Error Components							
S2SLS	0.004	0.034	0.029	0.001	0	0.001	0.033
	(0.015)	(0.135)	(0.107)	(0.022)	(0.063)	(0.057)	(0.049)
	[0.017]	[0.145]	[0.117]	[0.02]	[0.068]	[0.058]	
S3SLS	0.005	0.039	0.031	0.001	0.009	0.005	0.033
	(0.015)	(0.132)	(0.104)	(0.022)	(0.061)	(0.052)	(0.049)
	[0.017]	[0.147]	[0.116]	[0.021]	[0.067]	[0.056]	

[1] In each cell of columns 2-7, the upper number denotes the bias.
[2] The middle one in parentheses denotes the standard deviation.
[3] The bottom number denotes the RMSE.
[4] In each cell of column 8, the upper number denotes NOMAD.
[5] the lower one in parentheses denotes NORMSQD.

References

Allers, M. A. and Elhorst, J. P. (2011). A simultaneous equations model of fiscal policy interactions. *Journal of regional science*, 51(2):271–291.

Anselin, L. (1988). Model validation in spatial econometrics: a review and evaluation of alternative approaches. *International Regional Science Review*, 11(3):279–316.

Anselin, L. (2006). Spatial econometrics in: Mills tc, patterson k (eds) palgrave hand'book of econometrics: Volume 1, econometric theory.

Anselin, L. and Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics Textbooks and Monographs*, 155:237–290.

Balestra, P. and Varadharajan-Krishnakumar, J. (1987). Full information estimations of a system of simultaneous equations with error component structure. *Econometric Theory*, 3(02):223–246.

Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.

Baltagi, B. H. (1981). Simultaneous equations with error components. *Journal of econometrics*, 17(2):189–200.

Baltagi, B. H. (1984). A monte carlo study for pooling time series of cross-section data in the simultaneous equations model. *International Economic Review*, pages 603–624.

Baltagi, B. H. and Bresson, G. (2011). Maximum likelihood estimation and lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: An application to hedonic housing prices in paris. *Journal of Urban Economics*, 69(1):24–42.

Baltagi, B. H. and Deng, Y. (2015). Ec3sls estimator for a simultaneous system of spatial autoregressive equations with random effects. *Econometric Reviews*, 34(6-10):659–694.

Baltagi, B. H., Egger, P., and Pfaffermayr, M. (2013). A generalized spatial panel data model with random effects. *Econometric Reviews*, 32(5-6):650–685.

Baltagi, B. H. and Li, Q. (1992). A note on the estimation of simultaneous equations with error components. *Econometric Theory*, 8(1):113–119.

Baltagi, B. H. and Pirotte, A. (2011). Seemingly unrelated regressions with spatial error components. *Empirical Economics*, 40(1):5–49.

Barro, R. and Sala-i Martin, X. (1995). Economic growth. ed.

Cliff, A. D. and Ord, J. K. (1973). Spatial autocorrelation. Technical report, London: Pion.

Cliff, A. D. and Ord, J. K. (1981). *Spatial processes: models & applications*. Taylor & Francis.

Cornwell, C., Schmidt, P., and Wyhowski, D. (1992). Simultaneous equations and panel data. *Journal of Econometrics*, 51(1-2):151–181.

Cressie, N. (2015). *Statistics for spatial data*. John Wiley & Sons.

De Graaff, T., Van Oort, F. G., and Florax, R. J. (2012). Regional population–employment dynamics across different sectors of the economy. *Journal of Regional Science*, 52(1):60–84.

Doğan, O. and Taşpınar, S. (2013). Gmm estimation of spatial autoregressive models with moving average disturbances. *Regional Science and Urban Economics*, 43(6):903–926.

Fingleton, B. (2008). A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices. *Empirical Economics*, 34(1):35–57.

Gebremariam, G. H., Gebremedhin, T. G., and Schaeffer, P. V. (2011). Employment, income, and migration in appalachia: A spatial simultaneous equations approach. *Journal of Regional Science*, 51(1):102–120.

Goldsmith-Pinkham, P. and Imbens, G. W. (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics*, 31(3):253–264.

Haining, R. (1978). The moving average model for spatial interaction. *Transactions of the Institute of British Geographers*, pages 202–225.

Hauptmeier, S., Mittermaier, F., and Rincke, J. (2012). Fiscal competition over taxes and public inputs. *Regional science and urban economics*, 42(3):407–419.

Hepple, L. W. (2003). Bayesian and maximum likelihood estimation of the linear model with spatial moving average disturbances. *School of Geographical Sciences, University of Bristol, Working Papers Series, University of Bristol, School of Geographical Sciences.*

Hitiris, T. and Posnett, J. (1992). The determinants and effects of health expenditure in developed countries. *Journal of health economics*, 11(2):173–181.

Ho, C.-S. and Hite, D. (2008). The benefit of environmental improvement in the southeastern united states: Evidence from a simultaneous model of cancer mortality, toxic chemical releases and house values. *Papers in Regional Science*, 87(4):589–604.

Jeanty, P. W., Partridge, M., and Irwin, E. (2010). Estimation of a spatial simultaneous equation model of population migration and housing price dynamics. *Regional Science and Urban Economics*, 40(5):343–352.

Kapoor, M., Kelejian, H. H., and Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of econometrics*, 140(1):97–130.

Kelejian, H. H. and Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1):99–121.

Kelejian, H. H. and Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review*, 40(2):509–533.

Kelejian, H. H. and Prucha, I. R. (2004). Estimation of simultaneous systems of spatially interrelated cross sectional equations. *Journal of econometrics*, 118(1):27–50.

Kelejian, H. H., Prucha, I. R., and Yuzefovich, Y. (2004). Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: Large and small sample results. In James, P. L. and Kelley, R. P., editors, *Spatial and spatiotemporal econometrics*, volume 18, pages 163–198. Emerald Group Publishing Limited.

Kelejian, H. H. and Robinson, D. P. (1992). Spatial autocorrelation: A new computationally simple test with an application to per capita county police expenditures. *Regional Science and Urban Economics*, 22(3):317–331.

Lee, L.-f. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews*, 22(4):307–335.

LeSage, J. P. and Pace, R. K. (2009). *Introduction to Spatial Econometrics (Statistics, textbooks and monographs)*. CRC Press.

Liu, X. (2014). Identification and efficient estimation of simultaneous equations network models. *Journal of Business & Economic Statistics*, 32(4):516–536.

Liu, X. and Saraiva, P. (2017). Gmm estimation of spatial autoregressive models in a system of simultaneous equations with heteroskedasticity. *Econometric Reviews*, (just-accepted).

Lu, L. (2017). Simultaneous spatial panel data models with common shocks. *Working Paper*.

O'Connell, J. M. (1996). The relationship between health expenditures and the age structure of the population in oecd countries. *Health economics*, 5(6):573–578.

Parkin, D., McGuire, A., and Yule, B. (1987). Aggregate health care expenditures and national income: is health care a luxury good? *Journal of health economics*, 6(2):109–127.

Prucha, I. R. (1985). Maximum likelihood and instrumental variable estimation in simultaneous equation systems with error components. *International Economic Review*, pages 491–506.

Prucha, I. R., Drukker, D. M., and Egger, P. H. (2016). Simultaneous equations models with higher-order spatial or social network interactions. Technical report, Working paper, Department of Economics, University of Maryland. http://econweb. umd. edu/ prucha/papers/WP IRP PHE DMD 2016. pdf.

Sasser, W. E. (1969). *A finite-sample study of various simultaneous equation estimators...* Duke University.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, 70(1):65–94.

Wang, L., Li, K., and Wang, Z. (2014). Quasi maximum likelihood estimation for simultaneous spatial autoregressive models. *University Library of Munich, Germany.*

Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, pages 434–449.

Yang, K. and Lee, L.-f. (2017). Identification and qml estimation of multivariate and simultaneous equations spatial autoregressive models. *Journal of Econometrics*, 196(1):196–214.

Zenou, Y. (2017). Multivariate choices and identification of social interactions. *Journal of Applied Econometrics*, pages 1–37.