Parameter estimation of the spatial panel stochastic frontier model with random effects

Seya, Hajime

(Citation)
Applied Economics Letters,27(3):248-253

(Issue Date)
2020-02-06

(Resource Type)
journal article

(Version)
Accepted Manuscript

Parameter estimation of the spatial panel stochastic frontier model with random effects

Hajime Seya

Graduate School of Engineering Faculty of Engineering, Kobe University,

1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan,

E-mail: hseya@people.kobe-u.ac.jp

Abstract: This study has two objectives. First, it aims to construct a spatial panel stochastic frontier model with random effects based on the Cliff and Ord-type spatial panel model. Second, it proposes a parameter estimation method for this model. The method is based on a four-step procedure outlined by Kumbhakar and Heshmati (1995), which was modified by incorporating spatial autocorrelation.

Keywords: stochastic frontier model, spatial panel, random effects, generalized spatial two-stage least squares

JEL: C40; C50; C13

Acknowledgement

This study was funded by JSPS KAKENHI Grant Numbers 17K14738 and 18H03628.

I. Introduction

This study has two objectives. First, it aims to construct a spatial panel stochastic frontier model with random effects based on the Cliff and Ord-type spatial panel model. Although recently many relevant models have been proposed (e.g., Glass et al., 2016; Vidoli et al., 2016; Carvalho, 2018), we are not aware of any that introduce both the time variant and persistent (in)efficiencies. Second, we propose a parameter estimation method for this model. Our model, which is based on the four-step procedure outlined by Kumbhakar and Heshmati (1995), has an advantage in terms of computational burden.

II. Parameter estimation of panel stochastic frontier model with random effects

Stochastic frontier models can be defined in terms of their production function or cost function. We use cost function to illustrate our model. Consider the following cost function:

$$C_i = C(y_i, \varpi_i), \tag{1}$$

where y_i (i=1, …, N) denotes output for firm i and ϖ_i denotes factor price. Suppose that y_i is approximated by a $(K-1)\times 1$ attributes vector $\boldsymbol{s}_i$ and the $K\times 1$ vector is defined as $\boldsymbol{X}_i = [\boldsymbol{s}_i', \varpi_i]'$, then, the stochastic frontier cost function is given as:

$$C_i = C(\boldsymbol{X}_i)\cdot \exp(v_i)\cdot (CIE)_i, \tag{2}$$

where v_i represents the independent and identically distributed (i.i.d.) error term, and cost inefficiency (CIE) represents technical inefficiency. A CIE value of 1 indicates that technical inefficiency does not exist, whereas a value greater than 1 implies that it does. In our empirical analysis, we usually specify CIE for the ith observation as $CIE_i = \exp(u_i)$, where $u_i \geq 0$. If we simply assume a linear functional form for $C(\boldsymbol{X}_i)$ and consider the natural logarithm of both sides in Eq. (2), we have:

$$\ln C_i = \beta_0 + \sum_{k=1}^{K} \beta_k \ln X_{ki} + u_i + v_i. \tag{3}$$

One of the conventional assumptions on the distributions of v_i and u_i is the normal-half-normal model, where i.i.d. normal $N(0,\sigma_v^2)$ is assumed on v_i and i.i.d. half-normal $N^+(0,\sigma_u^2)$ is assumed on u_i.

Using panel data, the stochastic frontier cost function can be rewritten as:

$$\ln C_{i,t} = \beta_0 + \sum_{k=1}^{K} \beta_k \ln X_{ki,t} + \mu_i + u_{i,t} + v_{i,t}, \tag{4a}$$

where t (t = 1, …, T) denotes time. Various types of stochastic frontier models have been proposed using assumptions on μ_i, $u_{i,t}$, and $v_{i,t}$. In this paper, we follow the model and interpretation of Kumbhakar and Heshmati (1995), in which $v_{i,t}$ denotes the zero mean i.i.d. error term, and the exponential value of the sum of the time-invariant μ_i (≥ 0) and time-variant $u_{i,t}$ (≥ 0) "random effects" represents CIE, that is, $\exp(\mu_i + u_{i,t})$.

The parameters of Eq. (4a) may be estimated by a single stage maximum likelihood (ML) method (Colombi et al., 2014) or a multistage approach outlined in the study by Kumbhakar and Heshmati (1995) (see Smith and Wheat, 2012). The advantages of the latter approach include fewer distributional assumptions (as described later) and a lower computational burden, in exchange for loss of estimation efficiency. In this paper, we adopt the latter, more practical approach, giving greater weight to computational advantage.

In their study, Kumbhakar and Heshmati (1995) propose the following four-step parameter estimation method:

- STEP 1: Apart from the intercept, the regression coefficients β_k (k = 1, …, K) are estimated using the feasible generalized-least-squares (FGLS) method.
- STEP 2: The time-invariant term μ_i is estimated.
- STEP 3: The intercept β_0 and variances for $u_{i,t}$ and $v_{i,t}$ are estimated.
- STEP 4: The time-variant term $u_{i,t}$ is estimated.

Distribution assumptions [$v_{i,t}$~i.i.d. $N(0, \sigma_v^2)$ and $u_{i,t}$ ~i.i.d. $N^+(0, \sigma_u^2)$] are required only in Steps 3 and 4 and not in Steps 1 and 2.

Because we have assumed that $u_{i,t}$ and $v_{i,t}$ are nonnegative, their expectations are typically not equal to zero. Hence, we first transform Eq. (4a) as:

$$c_{i,t} = \beta_0^* + \sum_{k=1}^{K} \beta_k x_{ki,t} + \mu_i^* + \omega_{i,t}^* , \tag{4b}$$

where $\mu_i^* = \mu_i - E(\mu_i)$; $\omega_{i,t}^* = u_{i,t} - E(u_{i,t}) + v_{i,t}$; $\beta_0^* = \beta_0 + E(\mu_i) + E(u_{i,t})$, with $c_{i,t}$ = ln$C_{i,t}$ and $x_{ki,t}$ = ln$X_{ki,t}$. Because $\mu_i^*$and $\omega_{i,t}^*$ have zero expectations and homogeneous variances, we can apply the standard approach for the parameter estimation of the panel data model to Eq. (4b). By stacking Eq. (4b) with respect to i, we have:

$$\boldsymbol{c}_t = \beta_0^* \boldsymbol{\iota}_N + \boldsymbol{x}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t , \tag{4c}$$

where $\boldsymbol{c}_t$ denotes an N×1 vector consisting of $c_{i,t}$; $\boldsymbol{\iota}_m$ denotes an m×1 vector ones for an integer m (Here, m equals to N); $\boldsymbol{x}_t$ denotes an N×K matrix whose elements are given by $x_{ki,t}$; and $\boldsymbol{\beta}$ denotes a K×1 vector whose elements are given by β_k. Moreover, let $\boldsymbol{\varepsilon}_t = \boldsymbol{\mu}^* + \boldsymbol{\omega}_t^*$, where $\boldsymbol{\mu}^*$and $\boldsymbol{\omega}_t^*$ denote an N×1 vector whose elements are given by μ_i^* and $\omega_{i,t}^*$, respectively. By stacking Eq. (4c) with respect to t, we have:

$$\boldsymbol{c} = \beta_0^* \boldsymbol{\iota} + \boldsymbol{x}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \equiv \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} , \tag{4d}$$

where $\boldsymbol{Z} = (\boldsymbol{\iota}, \boldsymbol{x}),\ \boldsymbol{\gamma}' = (\beta_0^*, \boldsymbol{\beta}')$. The four-step procedure is detailed as follows:

STEP 1

In the first step, the standard random effect regression model for panel model is applied. The FGLS estimator of $\boldsymbol{\gamma}$, that is, $\hat{\boldsymbol{\gamma}}_{GLS}$, is given by:

$$\hat{\boldsymbol{\gamma}}_{GLS} = (\boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{c} , \tag{5}$$

where $\boldsymbol{V}$ is the variance–covariance matrix that satisfies:

$$\boldsymbol{V} = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma_\mu^2(\boldsymbol{\iota}_T\boldsymbol{\iota}_T' \otimes \boldsymbol{I}_N) + \sigma_\omega^2 \boldsymbol{I}_{NT} , \tag{6}$$

where $\boldsymbol{I}_m$ denotes the m×m identity matrix and $\otimes$ denotes the Kronecker product. Consistent estimators of variance parameters are given by:

$$\hat{\sigma}_\omega^2 = \frac{1}{NT}\sum_{i}^{N}\sum_{t}^{T}\left\{\left(c_{i,t} - \bar{c}_i\right) - \hat{\boldsymbol{\gamma}}_{OLS}'(\boldsymbol{x}_{i,t} - \bar{\boldsymbol{x}}_i)\right\}^2 , \tag{7a}$$

$$\hat{\sigma}_{\mu}^{2} = \frac{1}{N}\sum_{i}^{N}\left\{ (\bar{c}_i - \hat{\gamma}'_{OLS}\bar{\mathbf{x}}_i)^2 - \frac{1}{T}\hat{\sigma}_{\omega}^{2} \right\}^2 . \quad (7b)$$

where $\bar{c}_i, \bar{\mathbf{x}}_i$ denotes the means of $c_{i,t}$ and $\mathbf{x}_{i,t}$ over time and $\hat{\gamma}_{OLS}$ denotes the OLS estimate of $\boldsymbol{\gamma}$.

STEP 2

Using $\hat{\boldsymbol{\beta}}$, which is a subset of $\hat{\gamma}$, we can calculate the residual as $e_{i,t} = c_{i,t} - \hat{\boldsymbol{\beta}}'\mathbf{x}_{i,t}$, where $\mathbf{x}_{i,t}$ denotes the explanatory variables' vector for the *i*th observation. The residual $e_{i,t}$ includes information about $\beta_0, \mu_i, \omega_{i,t}$. Thus, the estimate of μ_i may be given relative to the best efficiency value as:

$$\hat{\mu}_i = \bar{e}_i - \min\{\bar{e}_i\} , \quad (8)$$

where $\bar{e}_i$ denotes the mean (over time) of $e_{i,t}$.

STEP 3

Using $\hat{\mu}_i$, we can calculate the residual $\varsigma_{i,t} = e_{i,t} - \hat{\mu}_i$, which includes information about $\beta_0 + \omega_{i,t} = \beta_0 + u_{i,t} + v_{i,t}$. In this step, we set the distribution assumptions on $v_{i,t}$ and $u_{i,t}$ as $v_{i,t}$~*i.i.d.* $N(0, \sigma_v^2)$ and $u_{i,t}$ ~*i.i.d.* $N^+(0, \sigma_u^2)$, respectively. Then, the log-likelihood function for observation *i* is given as:

$$\ln L(\beta_0, \sigma, \psi) = constant - \ln\sigma + \ln\Phi\left(\frac{r_{i,t}\psi}{\sigma}\right) - \frac{1}{2}\left(\frac{r_{i,t}}{\sigma}\right)^2 , \quad (9)$$

where $\psi = \sigma_u / \sigma_v$; $\sigma^2 = \sigma_v^2 + \sigma_u^2$; and $r_{i,t} = \varsigma_{i,t} - \beta_0$ and Φ denotes the cumulative distribution function of a standard normal distribution.

STEP 4

The last step is the estimation of $u_{i,t}$. Conventionally, following Jondrow et al. (1982), it may be estimated as:

$$\hat{u}_{i,t} = E(u_{i,t} \mid r_{i,t}) = \frac{\hat{\psi}\hat{\sigma}}{(1+\hat{\psi}^2)}\left[\frac{\phi(\hat{\psi}r_{i,t}/\hat{\sigma})}{1-\Phi(\hat{\psi}r_{i,t}/\hat{\sigma})} - \left(\frac{\hat{\psi}r_{i,t}}{\hat{\sigma}}\right)\right] , \quad (10)$$

where ϕ is the probability density function of a standard normal distribution. Once $\hat{u}_{i,t}$ is estimated, we can estimate *CIE* as $\exp(\hat{\mu}_i + \hat{u}_{i,t})$.

III. Parameter estimation of spatial panel stochastic frontier model with random effects

The spatial panel stochastic frontier model with random effects can be written as:

$$c_{i,t} = \beta_0 + \rho\sum_{j=1}^{N} w_{ij}c_{j,t} + \sum_{k=1}^{K}\beta_k x_{ki,t} + \varepsilon_{i,t} , \quad (11)$$

$$\varepsilon_{i,t} = \lambda\sum_{j=1}^{N} w_{ij}\varepsilon_{j,t} + \xi_{i,t}, \quad \xi_{i,t} = \mu_i + u_{i,t} + v_{i,t} , \quad (12a)$$

where ρ and λ denote the parameters representing the degrees of spatial autocorrelation in the dependent variables and errors, respectively (Kapoor et al., 2007). Following Kumbhakar and Heshmati's (1995) study and the parameter estimation method for the spatial panel model, we propose a parameter estimation method for the spatial panel stochastic frontier model with random effects.

First, we rewrite Eq. (12a) as

$$c_{i,t} = \beta_0^* + \rho \sum_{j=1}^{N} w_{ij} c_{j,t} + \sum_{k=1}^{K} \beta_k x_{ki,t} + \varepsilon_{i,t} ,$$

$$\varepsilon_{i,t} = \lambda \sum_{j=1}^{N} w_{ij} \varepsilon_{j,t} + \xi_{i,t}, \quad \xi_{i,t} = \mu_i^* + \omega_{i,t}^* , \tag{12b}$$

where $\mu_i^* = \mu_i - E(\mu_i)$; $\omega_{i,t}^* = u_{i,t} - E(u_{i,t}) + v_{i,t}$; $\beta_0^* = \beta_0 + E(\mu_i) + E(u_{i,t})$. Furthermore, we assume that $\mu_i^* \sim i.i.d.(0, \sigma_\mu^2)$ and $\omega_{i,t}^* \sim i.i.d.(0, \sigma_\omega^2)$. By stacking Eq. (12b) with respect to *i*, we have:

$$\boldsymbol{c}_t = \beta_0^* \boldsymbol{\iota}_N + \rho \boldsymbol{W} \boldsymbol{c}_t + \boldsymbol{x}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t ,$$

$$\boldsymbol{\varepsilon}_t = \lambda \boldsymbol{W} \boldsymbol{\varepsilon}_t + \boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t = \boldsymbol{\mu}^* + \boldsymbol{\omega}_t^* , \tag{12c}$$

where the $N \times N$ matrix $\boldsymbol{W}$ denotes a spatial weight matrix whose elements are given by w_{ij}. We assume that $\boldsymbol{W}$ is exogenously given and row-standardized for simplicity. By stacking Eq. (12c) with respect to *t*, we have:

$$\boldsymbol{c} = \beta_0^* \boldsymbol{\iota} + \rho (\boldsymbol{I}_T \otimes \boldsymbol{W}) \boldsymbol{c} + \boldsymbol{x} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \equiv \boldsymbol{Z} \boldsymbol{\gamma} + \boldsymbol{\varepsilon} ,$$

$$\boldsymbol{\varepsilon} = \lambda (\boldsymbol{I}_T \otimes \boldsymbol{W}) \boldsymbol{\varepsilon} + \boldsymbol{\xi} , \quad \boldsymbol{\xi} = (\boldsymbol{\iota}_T \otimes \boldsymbol{I}_N) \boldsymbol{\mu}^* + \boldsymbol{\omega}^* , \tag{12d}$$

where $\boldsymbol{Z} = (\boldsymbol{\iota}, (\boldsymbol{I}_T \otimes \boldsymbol{W}) \boldsymbol{c}, \boldsymbol{x})$; $\boldsymbol{\gamma}' = (\beta_0^*, \rho, \boldsymbol{\beta}')$. The variance–covariance matrices are given by:

$$\boldsymbol{\Omega}_\varepsilon = [\boldsymbol{I}_T \otimes (\boldsymbol{I}_N - \lambda W)^{-1}] \boldsymbol{\Omega}_\xi [\boldsymbol{I}_T \otimes (\boldsymbol{I}_N - \lambda W')^{-1}] , \tag{13a}$$

$$\boldsymbol{\Omega}_\xi = E(\boldsymbol{\xi} \boldsymbol{\xi}') = \sigma_\mu^2 (\boldsymbol{\iota}_T \boldsymbol{\iota}_T' \otimes \boldsymbol{I}_N) + \sigma_\omega^2 \boldsymbol{I}_{NT} . \tag{13b}$$

Here, based on the study by Baltagi (2008), we define the following two transformation matrices:

$$\boldsymbol{Q}_0 = \left(\boldsymbol{I}_T - \frac{\boldsymbol{\iota}_T \boldsymbol{\iota}_T'}{T} \right) \otimes \boldsymbol{I}_N , \quad \boldsymbol{Q}_1 = \frac{\boldsymbol{\iota}_T \boldsymbol{\iota}_T'}{T} \otimes \boldsymbol{I}_N . \tag{14}$$

The inverse of the variance-covariance matrix can then be simply written as $\boldsymbol{\Omega}_\xi^{-1} = \sigma_\omega^{-2} \boldsymbol{Q}_0 + \sigma_1^{-2} \boldsymbol{Q}_1$, where $\sigma_1^2 = \sigma_\omega^2 + T \sigma_\mu^2$. The remainder of the four-step procedure is detailed as follows:

STEP 1

In Step 1, using the generalized spatial two-stage least-squares (GS2SLS) method proposed by Kapoor et al. (2007), σ_ω^2, σ_1^2, and γ are estimated. First, under the no spatial error autocorrelation (λ = 0) and no persistent inefficiency ($\mu_i = 0$) condition, we estimate γ using the S2SLS method with an instrumental variable such as $\boldsymbol{H} = [\boldsymbol{x}, (\boldsymbol{I}_T \otimes \boldsymbol{W}), \boldsymbol{x}(\boldsymbol{I}_T \otimes \boldsymbol{W}^2)\boldsymbol{x}]$. Based on the S2SLS estimate of γ, that is, $\hat{\gamma}$, we can calculate the residual, $\hat{\boldsymbol{\varepsilon}} = \boldsymbol{c} - \boldsymbol{Z}\hat{\boldsymbol{\gamma}}$. Kapoor et al. (2007) defined the following moment conditions for the GMM estimators of the spatial error autocorrelation parameter λ and two variance parameters (σ_ω^2 and σ_1^2).

$$E\begin{bmatrix} \frac{1}{N(T-1)}\boldsymbol{\xi}'\boldsymbol{Q}_0\boldsymbol{\xi} \\ \frac{1}{N(T-1)}\breve{\boldsymbol{\xi}}'\boldsymbol{Q}_0\breve{\boldsymbol{\xi}} \\ \frac{1}{N(T-1)}\breve{\boldsymbol{\xi}}'\boldsymbol{Q}_0\boldsymbol{\xi} \\ \frac{1}{N}\boldsymbol{\xi}'\boldsymbol{Q}_1\boldsymbol{\xi} \\ \frac{1}{N}\breve{\boldsymbol{\xi}}'\boldsymbol{Q}_1\breve{\boldsymbol{\xi}} \\ \frac{1}{N}\breve{\boldsymbol{\xi}}'\boldsymbol{Q}_1\boldsymbol{\xi} \end{bmatrix} = \begin{bmatrix} \sigma_\omega^2 \\ \sigma_\omega^2 \frac{1}{N} tr(\boldsymbol{W}'\boldsymbol{W}) \\ 0 \\ \sigma_1^2 \\ \sigma_1^2 \frac{1}{N} tr(\boldsymbol{W}'\boldsymbol{W}) \\ 0 \end{bmatrix}, \tag{15}$$

where $\boldsymbol{\xi} = \boldsymbol{\varepsilon} - \lambda\breve{\boldsymbol{\varepsilon}}$; $\breve{\boldsymbol{\xi}} = \breve{\boldsymbol{\varepsilon}} - \lambda\breve{\breve{\boldsymbol{\varepsilon}}}$; $\breve{\boldsymbol{\varepsilon}} = (\boldsymbol{I}_T \otimes \boldsymbol{W})\boldsymbol{\varepsilon}$; and $\breve{\breve{\boldsymbol{\varepsilon}}} = (\boldsymbol{I}_T \otimes \boldsymbol{W})\breve{\boldsymbol{\varepsilon}}$. Sample analogues to Eq. (15), where the expectation operator is omitted, can be written as:

$$\begin{bmatrix} \frac{1}{N(T-1)}\tilde{\boldsymbol{\xi}}'\boldsymbol{Q}_0\tilde{\boldsymbol{\xi}} \\ \frac{1}{N(T-1)}\tilde{\breve{\boldsymbol{\xi}}}'\boldsymbol{Q}_0\tilde{\breve{\boldsymbol{\xi}}} \\ \frac{1}{N(T-1)}\tilde{\breve{\boldsymbol{\xi}}}'\boldsymbol{Q}_0\tilde{\boldsymbol{\xi}} \\ \frac{1}{N}\tilde{\boldsymbol{\xi}}'\boldsymbol{Q}_1\tilde{\boldsymbol{\xi}} \\ \frac{1}{N}\tilde{\breve{\boldsymbol{\xi}}}'\boldsymbol{Q}_1\tilde{\breve{\boldsymbol{\xi}}} \\ \frac{1}{N}\tilde{\breve{\boldsymbol{\xi}}}'\boldsymbol{Q}_1\tilde{\boldsymbol{\xi}} \end{bmatrix} = \begin{bmatrix} \sigma_\omega^2 \\ \sigma_\omega^2 \frac{1}{N} tr(\boldsymbol{W}'\boldsymbol{W}) \\ 0 \\ \sigma_1^2 \\ \sigma_1^2 \frac{1}{N} tr(\boldsymbol{W}'\boldsymbol{W}) \\ 0 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{bmatrix}. \tag{16}$$

Based on the first three conditions in Eq. (16), the parameters λ and σ_ω^2 can be estimated by applying the nonlinear least-squares (NLS) method, assigning equal weights to each condition as:

$$(\hat{\lambda}, \hat{\sigma}_\omega^2) = \arg\min_{\lambda, \sigma_\omega^2}[\delta_1^2 + \delta_2^2 + \delta_3^2]. \tag{17}$$

Using the estimates of $\hat{\lambda}, \hat{\sigma}_\omega^2$ obtained from Eq. (17), the σ_1^2 estimate can be obtained using the fourth condition in Eq. (16) as

$$\hat{\sigma}_1^2 = \frac{1}{N}\tilde{\boldsymbol{\xi}}'\boldsymbol{Q}_1\tilde{\boldsymbol{\xi}}. \tag{18}$$

Using the variance estimates obtained from Eqs. (17) and (18) as initial values, more efficient GMM estimates can be calculated by minimizing:

$$(\hat{\hat{\lambda}}, \hat{\hat{\sigma}}_\omega^2) = \arg\min_{\lambda, \sigma_\omega^2}[\boldsymbol{\delta}'\boldsymbol{R}^{-1}\boldsymbol{\delta}], \tag{19}$$

where

$$\boldsymbol{R}=\begin{bmatrix}\frac{1}{T-1}\sigma_\omega^4 & 0\\ 0 & \sigma_1^4\end{bmatrix}\otimes\boldsymbol{T}_W\ ;\ \ \boldsymbol{T}_W=\begin{bmatrix}2 & 2tr\left(\frac{\boldsymbol{W'W}}{N}\right) & 0\\ 2tr\left(\frac{\boldsymbol{W'W}}{N}\right) & 2tr\left(\frac{\boldsymbol{W'WW'W}}{N}\right) & tr\left(\frac{\boldsymbol{W'W}(\boldsymbol{W'}+\boldsymbol{W})}{N}\right)\\ 0 & tr\left(\frac{\boldsymbol{W'W}(\boldsymbol{W'}+\boldsymbol{W})}{N}\right) & tr\left(\frac{\boldsymbol{WW}+\boldsymbol{W'W}}{N}\right)\end{bmatrix}. \tag{20}$$

After obtaining the spatial error parameter estimate, a transformation is conducted to remove spatial error autocorrelation as $\boldsymbol{c}^*=[\boldsymbol{I}_T\otimes(\boldsymbol{I}_N-\hat{\hat{\lambda}}\boldsymbol{W})]\boldsymbol{c}$; $\boldsymbol{Z}^*=[\boldsymbol{I}_T\otimes(\boldsymbol{I}_N-\hat{\hat{\lambda}}\boldsymbol{W})]\boldsymbol{Z}$; $\boldsymbol{\xi}=[\boldsymbol{I}_T\otimes(\boldsymbol{I}_N-\hat{\hat{\lambda}}\boldsymbol{W})]\boldsymbol{\varepsilon}$. Then, we obtain

$$\boldsymbol{c}^*=\boldsymbol{Z}^*\boldsymbol{\gamma}+\boldsymbol{\xi}\ , \tag{21}$$

$$\hat{\boldsymbol{\Omega}}_\xi=\hat{\hat{\sigma}}_\omega^2\boldsymbol{Q}_0+\hat{\sigma}_1^2\boldsymbol{Q}_1\ .$$

Here, because $\boldsymbol{Z}^*$ includes an endogenous variable, that is, the spatial lag of the dependent variable, the estimation of $\boldsymbol{\gamma}$ depends on a S2SLS method that differs from the that in the case presented by Kapoor et al. (2007). Baltagi and Liu (2011) propose four different S2SLS estimators: [1] a random-effects model (RE-S2SLS), [2] a fixed-effects model (FE-S2SLS), [3] a between-effects model (BE-S2SLS), and [4] an error component model (EC-S2SLS). We can summarize these estimators as follows:

[1] RE-S2SLS estimator:

$$\hat{\gamma}_{RE-S2SLS}=[\dot{\boldsymbol{Z}}'^*\boldsymbol{P}_{\dot{\boldsymbol{H}}}\dot{\boldsymbol{Z}}^*]^{-1}\dot{\boldsymbol{Z}}'^*\boldsymbol{P}_{\dot{\boldsymbol{H}}}\dot{\boldsymbol{c}}^*\ , \tag{22}$$

$$\text{var}(\hat{\gamma}_{RE-S2SLS})=[\dot{\boldsymbol{Z}}'^*\boldsymbol{P}_{\dot{\boldsymbol{H}}}\dot{\boldsymbol{Z}}^*]^{-1^*}\ ,$$

where $\boldsymbol{P}_{\dot{\boldsymbol{H}}}=\dot{\boldsymbol{H}}(\dot{\boldsymbol{H}}'\dot{\boldsymbol{H}})^{-1}\dot{\boldsymbol{H}}'$, and the instrument variable can be defined as:

$$\dot{\boldsymbol{H}}=[\hat{\boldsymbol{\Omega}}_\xi^{-1/2}\boldsymbol{x},\hat{\boldsymbol{\Omega}}_\xi^{-1/2}(\boldsymbol{I}_T\otimes\boldsymbol{W})\boldsymbol{x},\hat{\boldsymbol{\Omega}}_\xi^{-1/2}(\boldsymbol{I}_T\otimes\boldsymbol{W}^2)\boldsymbol{x}]=[\dot{\boldsymbol{x}},(\boldsymbol{I}_T\otimes\boldsymbol{W})\dot{\boldsymbol{x}},(\boldsymbol{I}_T\otimes\boldsymbol{W}^2)\dot{\boldsymbol{x}}]\ , \tag{23}$$

where $\dot{\boldsymbol{c}}^*=\dot{\boldsymbol{Z}}^*\boldsymbol{\gamma}+\dot{\boldsymbol{\xi}}$ ($\dot{\boldsymbol{Z}}^*=\hat{\boldsymbol{\Omega}}_\xi^{-1/2}\boldsymbol{Z}^*$; $\dot{\boldsymbol{c}}^*=\hat{\boldsymbol{\Omega}}_\xi^{-1/2}\boldsymbol{c}^*$; $\dot{\boldsymbol{\xi}}=\hat{\boldsymbol{\Omega}}_\xi^{-1/2}\boldsymbol{\xi}$ with $\hat{\boldsymbol{\Omega}}_\xi^{-1/2}=(\hat{\sigma}_\omega^{-1}\boldsymbol{Q}_0+\hat{\sigma}_1^{-1}\boldsymbol{Q}_1)$).

[2] FE-S2SLS estimator:

$$\hat{\gamma}_{FE-S2SLS}=[\ddot{\boldsymbol{Z}}'^*\boldsymbol{P}_{\ddot{\boldsymbol{H}}}\ddot{\boldsymbol{Z}}^*]^{-1}\ddot{\boldsymbol{Z}}'^*\boldsymbol{P}_{\ddot{\boldsymbol{H}}}\ddot{\boldsymbol{c}}^*\ , \tag{24}$$

where $\boldsymbol{P}_{\ddot{\boldsymbol{H}}}=\ddot{\boldsymbol{H}}(\ddot{\boldsymbol{H}}'\ddot{\boldsymbol{H}})^{-1}\ddot{\boldsymbol{H}}'$, and the instrument variable can be defined as:

$$\ddot{\boldsymbol{H}}=[\boldsymbol{Q}_0\boldsymbol{x},\boldsymbol{Q}_0(\boldsymbol{I}_T\otimes\boldsymbol{W})\boldsymbol{x},\boldsymbol{Q}_0(\boldsymbol{I}_T\otimes\boldsymbol{W}^2)\boldsymbol{x}]=[\ddot{\boldsymbol{x}},(\boldsymbol{I}_T\otimes\boldsymbol{W})\ddot{\boldsymbol{x}},(\boldsymbol{I}_T\otimes\boldsymbol{W}^2)\ddot{\boldsymbol{x}}]\ , \tag{25}$$

where $\ddot{\boldsymbol{c}}^*=\ddot{\boldsymbol{Z}}^*\boldsymbol{\gamma}+\ddot{\boldsymbol{\xi}}$ ($\ddot{\boldsymbol{Z}}^*=\boldsymbol{Q}_0\boldsymbol{Z}^*$; $\ddot{\boldsymbol{c}}^*=\boldsymbol{Q}_0\boldsymbol{c}^*$; $\ddot{\boldsymbol{\xi}}=\boldsymbol{Q}_0\boldsymbol{\xi}$).

[3] BE-S2SLS estimator:

$$\hat{\gamma}_{BE-S2SLS} = [\dddot{\boldsymbol{Z}}'^{*} \boldsymbol{P}_{\dddot{\boldsymbol{H}}} \dddot{\boldsymbol{Z}}^{*}]^{-1} \dddot{\boldsymbol{Z}}'^{*} \boldsymbol{P}_{\dddot{\boldsymbol{H}}} \dddot{\boldsymbol{c}}^{*}, \tag{26}$$

where $\boldsymbol{P}_{\dddot{\boldsymbol{H}}} = \dddot{\boldsymbol{H}}(\dddot{\boldsymbol{H}}'\dddot{\boldsymbol{H}})^{-1}\dddot{\boldsymbol{H}}'$, and the instrument variable can be defined as:

$$\dddot{\boldsymbol{H}} = [\boldsymbol{Q}_1\boldsymbol{x}, \boldsymbol{Q}_1(\boldsymbol{I}_T \otimes \boldsymbol{W})\boldsymbol{x}, \boldsymbol{Q}_1(\boldsymbol{I}_T \otimes \boldsymbol{W}^2)\boldsymbol{x}] = [\dddot{\boldsymbol{x}}, (\boldsymbol{I}_T \otimes \boldsymbol{W})\dddot{\boldsymbol{x}}, (\boldsymbol{I}_T \otimes \boldsymbol{W}^2)\dddot{\boldsymbol{x}}], \tag{27}$$

where $\dddot{\boldsymbol{c}}^{*} = \dddot{\boldsymbol{Z}}^{*}\gamma + \dddot{\boldsymbol{\xi}}$ ($\dddot{\boldsymbol{Z}}^{*} = \boldsymbol{Q}_1\boldsymbol{Z}^{*}$; $\dddot{\boldsymbol{c}}^{*} = \boldsymbol{Q}_1\boldsymbol{c}^{*}$; $\dddot{\boldsymbol{\xi}} = \boldsymbol{Q}_1\boldsymbol{\xi}$).

[4] EC-S2SLS estimator:

The EC-S2SLS estimator can be expressed as the weighted combination of the FE-S2SLS estimator and the BE-S2SLS estimator (Baltagi and Liu, 2011), which is given as:

$$\hat{\gamma}_{EC-S2SLS} = \left[\frac{\ddot{\boldsymbol{Z}}'^{*} \boldsymbol{P}_{\ddot{\boldsymbol{H}}} \ddot{\boldsymbol{Z}}^{*}}{\hat{\sigma}_{\omega}^2} + \frac{\dddot{\boldsymbol{Z}}'^{*} \boldsymbol{P}_{\dddot{\boldsymbol{H}}} \dddot{\boldsymbol{Z}}^{*}}{\hat{\sigma}_{1}^2} \right]^{-1} \left[\frac{\ddot{\boldsymbol{Z}}'^{*} \boldsymbol{P}_{\ddot{\boldsymbol{H}}} \ddot{\boldsymbol{c}}^{*}}{\hat{\sigma}_{\omega}^2} + \frac{\dddot{\boldsymbol{Z}}'^{*} \boldsymbol{P}_{\dddot{\boldsymbol{H}}} \dddot{\boldsymbol{c}}^{*}}{\hat{\sigma}_{1}^2} \right], \tag{28}$$

$$\mathrm{var}(\hat{\gamma}_{EC-S2SLS}) = \left[\frac{\ddot{\boldsymbol{Z}}'^{*} \boldsymbol{P}_{\ddot{\boldsymbol{H}}} \ddot{\boldsymbol{Z}}^{*}}{\hat{\sigma}_{\omega}^2} + \frac{\dddot{\boldsymbol{Z}}'^{*} \boldsymbol{P}_{\dddot{\boldsymbol{H}}} \dddot{\boldsymbol{Z}}^{*}}{\hat{\sigma}_{1}^2} \right]^{-1}.$$

The EC-S2SLS estimator can be obtained by applying the S2SLS with $\ddddot{\boldsymbol{H}} = [\ddot{\boldsymbol{H}}, \dddot{\boldsymbol{H}}]$ as instrument to $\dot{\boldsymbol{c}}^{*} = \dot{\boldsymbol{Z}}^{*}\gamma + \dot{\boldsymbol{\xi}}$ (see Eq. (23)). The empirical difficulty in using the FE-S2SLS is that time-invariant explanatory variables are extracted by multiplying them by $\boldsymbol{Q}_0$; thus, we cannot obtain a consistent estimator for $\boldsymbol{Q}_1\boldsymbol{\xi}$, which is required in the GMM procedure. Hence, following Piras's (2013) approach, we use the following estimation procedure: First, by using the FE-S2SLS estimation instrumented using $\ddot{\boldsymbol{H}}$, we obtain the estimate for $\boldsymbol{Q}_0\boldsymbol{\xi}$ and the residual $\ddot{\boldsymbol{c}}^{*} - \ddot{\boldsymbol{Z}}^{*}\hat{\gamma}_{FE-S2SLS}$. Subsequently, using the BE-S2SLS estimation instrumented using $\dddot{\boldsymbol{H}}$, we obtain the estimate for $\boldsymbol{Q}_1\boldsymbol{\xi}$ and the residual $\dddot{\boldsymbol{c}}^{*} - \dddot{\boldsymbol{Z}}^{*}\hat{\gamma}_{BE-S2SLS}$. Then, using these residuals, we estimate the spatial parameter of the error term and two variance parameters with the GMM procedure. Finally, we apply the RE-S2SLS or EC-S2SLS using $\dot{\boldsymbol{H}}$ or $\ddddot{\boldsymbol{H}}$ as instruments.

STEP 2

Using, for instance, $\hat{\gamma}_{EC-S2SLS}$, we calculate the residual $e_{i,t} = c_{i,t} - \hat{\rho}\sum_{j=1}^{N} w_{ij} y_{j,t} - \hat{\boldsymbol{\beta}}'\boldsymbol{x}_{i,t} - \hat{\lambda}\sum_{j=1}^{N} w_{ij} e_{j,t}^{*}$, where $e_{i,t}^{*} = c_{i,t} - \hat{\beta}_0^{*} - \hat{\rho}\sum_{j=1}^{N} w_{ij} y_{j,t} - \hat{\boldsymbol{\beta}}'\boldsymbol{x}_{i,t}$. Then, the residual $e_{i,t}$ includes information about $\beta_0, \mu_i, \omega_{i,t}$. The estimate of μ_i may then be provided as relative to the best efficiency value as $\hat{\mu}_i = \bar{e}_i - \min\{\bar{e}_i\}$.

STEPS 3 and 4

Steps 3 and 4 are substantially identical to those followed by Kumbhakar and Heshmati (1995).

IV. Possible model extensions

In Eq.(4a) and (12a), we assumed that μ_i denotes a random effect. However, μ_i can actually be treated as either a fixed or random effect. If we treat μ_i to be a fixed effect, the RE-S2SLS may be replaced by within (i.e., FE-S2SLS) estimator. The non-spatial fixed effect version of the multistep procedure used by Kumbhakar and Heshmati is applied in the study by Adom et al. (2018). Thus, the consideration of a fixed effect case is possible in a straightforward manner.

Also, it may be worthwhile to mention that although Kumbhakar and Heshmati (1995) assume that (the exponential value of) the term μ_i can be interpreted as persistent inefficiency, it can also be interpreted as firm-specific effect, as standard panel model (Green, 2005), or in between (Kumbhakar et al., 2014). Thus, we need to admit that there is no unique 'correct' model, and model selection shall be context dependent. Nevertheless, it is also important to develop spatial models for an in-between case such as that highlighted by Kumbhakar et al. (2014).

References

1) Adom, P.K., Amakye, K., Abrokwa, K.K. and Quaidoo, C. 2018. Estimate of transient and persistent energy efficiency in Africa: A stochastic frontier approach. *Energy Conversion and Management*, 166, 556–568.

2) Baltagi, B.H. 2008. *Econometric Analysis of Panel Data, 4th Edition*. Wiley, New York.

3) Baltagi, B.H. and Liu, L. 2011. Instrumental variable estimation of a spatial autoregressive panel model with random effects. *Economics Letters*, 111 (2), 135–137.

4) Carvalho, A. 2018. Efficiency spillovers in Bayesian stochastic frontier models: Application to electricity distribution in New Zealand. *Spatial Economic Analysis*, 13(2), 171–190.

5) Colombi, R., Kumbhakar, S. C., Martini, G. and Vittadini, G. 2014. Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. *Journal of Productivity Analysis,* 42(2), 123–136.

6) Glass, A. J., Kenjegalieva, K. and Sickles, R.C. 2016. A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers. *Journal of Econometrics*, 190(2), 289–300.

7) Greene, W. 2005. Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis*, 23(1), 7–32.

8) Jondrow, J., Lovell, C.A.K., Materov, I.S. and Schmidt, P. 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19 (2–3), 233–238.

9) 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.

10) Kumbhakar, S.C. and Heshmati, A. 1995. Efficiency measurement in Swedish dairy farms: An application of rotating panel data, 1976–88. *American Journal of Agricultural Economics*, 77(3), 660–674.

11) Kumbhakar, S.C., Lien, G. and Hardaker, J.B. 2014. Technical efficiency in competing panel data models: A study of Norwegian grain farming. J*ournal of Productivity Analysis*, 41(2), 321–337.

12) Piras, G. 2013. Efficient GMM estimation of a Cliff and Ord panel data model with random effects. *Spatial Economic Analysis*, 8(3), 370–388.

13) Smith, A.S. and Wheat, P. 2012. Estimation of cost inefficiency in panel data models with firm specific and sub-company specific effects. *Journal of Productivity Analysis*, 37(1), 27–40.

14) Vidoli, F., Cardillo, C., Fusco, E. and Canello, J. 2016. Spatial nonstationarity in the stochastic frontier model:

An application to the Italian wine industry. *Regional Science and Urban Economics*, 61(C), 153–164.