Econometric Theory, **22**, 2006, 973–984. Printed in the United States of America.
DOI: 10.1017/S0266466606060464

NOTES AND PROBLEMS

RANDOM EFFECTS AND SPATIAL AUTOCORRELATION WITH EQUAL WEIGHTS

BADI H. BALTAGI
Syracuse University

This note considers a panel data regression model with spatial autoregressive disturbances and random effects where the weight matrix is normalized and has equal elements. This is motivated by Kelejian, Prucha, and Yuzefovich (2005, *Journal of Regional Science*, forthcoming), who argue that such a weighting matrix, having blocks of equal elements, might be considered when units are equally distant within certain neighborhoods but unrelated between neighborhoods. We derive a simple weighted least squares transformation that obtains generalized least squares (GLS) on this model as a simple ordinary least squares (OLS). For the special case of a spatial panel model with no random effects, we obtain two sufficient conditions where GLS on this model is equivalent to OLS. Finally, we show that these results, for the equal weight matrix, hold whether we use the spatial autoregressive specification, the spatial moving average specification, the spatial error components specification, or the Kapoor, Kelejian, and Prucha (2005, *Journal of Econometrics*, forthcoming) alternative to modeling panel data with spatially correlated error components.

1. INTRODUCTION

Spatial models deal with correlation across spatial units, usually in a cross-section setting. However, Anselin (1988) also considered spatial panel data models that allow for random effects across these units. See also Baltagi, Song, and Koh (2003) and Kapoor, Kelejian, and Prucha (2005) for two recent studies on testing and estimation in these models and Case (1991) and Holtz-Eakin (1994) for two empirical applications. This note focuses on a panel data regression model with spatial autoregressive disturbances and random effects where the weight matrix is normalized and has equal elements. This is motivated by

I thank Paolo Paruolo and an anonymous referee for helpful comments and suggestions. Address correspondence to Badi H. Baltagi, Department of Economics and Center for Policy Research, Syracuse University, 426 Eggers Hall, Syracuse, NY 13244-1020, USA; e-mail: bbaltagi@maxwell.syr.edu.

© 2006 Cambridge University Press 0266-4666/06 $12.00

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

Kelejian and Prucha (2002) and Kelejian, Prucha, and Yuzefovich (2005), who use this weight matrix in the context of a spatially lagged dependent variable model. Kelejian et al. (2005) argue that such a weighting matrix having blocks of equal elements might be considered when units are equally distant within certain neighborhoods but unrelated between neighborhoods. Section 2 introduces a panel regression model with spatial autoregressive (SAR) disturbances and random effects. We show that for the equal weight matrix case, one can derive a simple weighted least squares transformation that obtains generalized least squares (GLS) on this model as a simple ordinary least squares (OLS). For the special case of a spatial panel model with no random effects, we obtain two sufficient conditions where GLS on this model is equivalent to OLS. Section 3 extends these results to other types of spatial error specifications, including the spatial moving average (SMA) and spatial error components (SEC) cases and also an alternative panel data regression model with spatially correlated error components suggested by Kapoor et al. (2005).

2. THE MODEL

Consider the following panel data regression model:

$$y_{ti} = \alpha + X'_{ti}\beta + u_{ti}, \qquad i = 1, \ldots, N; \qquad t = 1, \ldots, T, \tag{1}$$

where y_{ti} is the observation on the ith region for the tth time period, X_{ti} denotes the $k \times 1$ vector of observations on the nonstochastic regressors and u_{ti} is the regression disturbance. Here α is a scalar and β is a $k \times 1$ vector of slope parameters. In vector form, the disturbance vector of (1) is assumed to have random region effects and SAR remainder disturbances:

$$u_t = \mu + \epsilon_t, \tag{2}$$

with

$$\epsilon_t = \lambda W_N \epsilon_t + \nu_t, \tag{3}$$

where $u'_t = (u_{t1}, \ldots, u_{tN})$ and ϵ_t and ν_t are similarly defined. Here $\mu' = (\mu_1, \mu_2, \ldots, \mu_N)$ denote the vector of random region effects, which are assumed to be $iid(0, \sigma^2_\mu)$, λ is the scalar SAR coefficient with $|\lambda| < 1$, whereas $\nu_{ti} \sim iid(0, \sigma^2_\nu)$. We assume that μ and ν are independent. The term W_N is an $N \times N$ weighting matrix with zero elements across the diagonal and *equal elements* $(1/(N-1))$ off the diagonal. In other words, the disturbance for each unit is related to an average of the $(N-1)$ disturbances of the remaining units. Such a weighting matrix was recently considered by Kelejian and Prucha (2002), Lee (2002), and Kelejian et al. (2005) and would naturally arise if all units are neighbors to each other and there is no other reasonable or observable measure of distance between them. Kelejian and Prucha (2002) consider the case of a

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

spatially lagged dependent variable model with a row normalized weighting matrix with equal elements. They show that OLS and two-stage least squares (2SLS) are *inconsistent* unless panel data are available. Kelejian et al. (2005) give exact small-sample results that corroborate the earlier asymptotic findings. In addition, they demonstrate that for the spatial panel data case with fixed effects across time, OLS and 2SLS are both inconsistent. Lee (2002), on the other hand, shows that OLS can be *consistent* in an economic spatial environment where each unit can be influenced aggregately by a significant portion of units in the population.

One can rewrite (3) as

$$\epsilon_t = (I_N - \lambda W_N)^{-1}\nu_t = B_N^{-1}\nu_t, \tag{4}$$

where $B_N = I_N - \lambda W_N$. The model (1) can be rewritten in matrix notation as

$$y = \alpha\iota_{NT} + X\beta + u = Z\gamma + u, \tag{5}$$

where y is of dimension $NT \times 1$, ι_{NT} is a vector of ones of dimension NT, X is $NT \times k$, u is $NT \times 1$, and $Z = (\iota_{NT}, X)$. Here X is assumed to be of full column rank, and its elements are assumed to be bounded in absolute value. The disturbance term can be written in vector form as

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes B_N^{-1})\nu, \tag{6}$$

where $\nu' = (\nu_1', \nu_2', \ldots, \nu_T')$ and u is similarly defined, ι_T is a vector of ones of dimension T and I_T is an identity matrix of dimension T, and $\otimes$ denotes the Kronecker product. Under these assumptions, the variance-covariance matrix of u can be written as

$$\Omega = \sigma_\mu^2(J_T \otimes I_N) + \sigma_\nu^2(I_T \otimes (B_N'B_N)^{-1}). \tag{7}$$

Kelejian and Prucha (2002) showed that W_N can be written as

$$W_N = \frac{J_N}{(N-1)} - \frac{I_N}{(N-1)}, \tag{8}$$

where J_N is a matrix of ones and I_N is an identity matrix, both of dimension N. Defining $\bar{J}_N = J_N/N$ and $E_N = I_N - \bar{J}_N$, and replacing J_N and I_N by their equivalent terms $N\bar{J}_N$ and $(E_N + \bar{J}_N)$ and collecting like terms (see Wansbeek and Kapteyn, 1982), one gets

$$W_N = \bar{J}_N - \frac{E_N}{(N-1)} \tag{9}$$

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

with

$$B_N = (I_N - \lambda W_N) = \frac{(N-1+\lambda)}{(N-1)} E_N + (1-\lambda)\bar{J}_N.$$

Therefore

$$B_N^{-1} = (I_N - \lambda W_N)^{-1} = \frac{(N-1)}{(N-1+\lambda)} E_N + \frac{1}{(1-\lambda)} \bar{J}_N. \tag{10}$$

Note that W_N is symmetric and

$$(B_N' B_N)^{-1} = \frac{(N-1)^2}{(N-1+\lambda)^2} E_N + \frac{1}{(1-\lambda)^2} \bar{J}_N = c_1 E_N + c_2 \bar{J}_N, \tag{11}$$

where $c_1 = (N-1)^2/(N-1+\lambda)^2$ and $c_2 = 1/(1-\lambda)^2$.

Replacing J_T and I_T by their equivalent terms $T\bar{J}_T$ and $(E_T + \bar{J}_T)$ in (7), one gets

$$\Omega = T\sigma_\mu^2(\bar{J}_T \otimes (E_N + \bar{J}_N)) + \sigma_\nu^2((E_T + \bar{J}_T) \otimes (c_1 E_N + c_2 \bar{J}_N)). \tag{12}$$

Collecting like terms, we obtain the spectral decomposition of Ω,

$$\begin{aligned}\Omega = {} & (T\sigma_\mu^2 + c_1\sigma_\nu^2)(\bar{J}_T \otimes E_N) + (T\sigma_\mu^2 + c_2\sigma_\nu^2)(\bar{J}_T \otimes \bar{J}_N) \\ & + c_1\sigma_\nu^2(E_T \otimes E_N) + c_2\sigma_\nu^2(E_T \otimes \bar{J}_N).\end{aligned} \tag{13}$$

Hence

$$\begin{aligned}\Omega^{-1} = {} & \frac{1}{(T\sigma_\mu^2 + c_1\sigma_\nu^2)} (\bar{J}_T \otimes E_N) + \frac{1}{(T\sigma_\mu^2 + c_2\sigma_\nu^2)} (\bar{J}_T \otimes \bar{J}_N) \\ & + \frac{1}{c_1\sigma_\nu^2} (E_T \otimes E_N) + \frac{1}{c_2\sigma_\nu^2} (E_T \otimes \bar{J}_N)\end{aligned} \tag{14}$$

and

$$\begin{aligned}\Omega^{-1/2} = {} & \frac{1}{\sqrt{(T\sigma_\mu^2 + c_1\sigma_\nu^2)}} (\bar{J}_T \otimes E_N) + \frac{1}{\sqrt{(T\sigma_\mu^2 + c_2\sigma_\nu^2)}} (\bar{J}_T \otimes \bar{J}_N) \\ & + \frac{1}{\sqrt{c_1\sigma_\nu^2}} (E_T \otimes E_N) + \frac{1}{\sqrt{c_2\sigma_\nu^2}} (E_T \otimes \bar{J}_N).\end{aligned} \tag{15}$$

Note that if $\lambda = 0$, so that there is no spatial autocorrelation, then $c_1 = c_2 = 1$ and Ω in (13) reduces to the familiar variance-covariance matrix of the random effects panel data model given by $\Omega = (T\sigma_\mu^2 + \sigma_\nu^2)(\bar{J}_T \otimes I_N) + \sigma_\nu^2(E_T \otimes I_N)$ (see Baltagi, 2005). Premultiplying the regression model in (5) by $\sigma_\nu\sqrt{c_1}\Omega^{-1/2}$

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

from (15), one gets a similar transformation to the one suggested by Fuller and Battese (1973) for the random effects panel data model. In fact $y^* = \sigma_\nu \sqrt{c_1}\Omega^{-1/2} y$ will have typical elements $y^*_{ti} = (y_{ti} - \theta_1 \bar{y}_{t.} - \theta_2 \bar{y}_{.i} + \theta_3 \bar{y}_{..})$ where $\bar{y}_{t.}$ denotes the sample average over regions, $\bar{y}_{.i}$ denotes the sample average over time, and $\bar{y}_{..}$ denotes the average over the entire sample. The θ's can be easily obtained from the corresponding $(c_1, c_2, \sigma^2_\mu, \sigma^2_\nu)$ parameters. For example, θ_1 is the coefficient corresponding to $(I_T \otimes \bar{J}_N)$, which can be verified to be $(1 - \sqrt{c_1/c_2})$, θ_2 is the coefficient corresponding to $(\bar{J}_T \otimes I_N)$, which can be verified to be $(1 - \sqrt{c_1 \sigma^2_\nu/(T\sigma^2_\mu + c_1\sigma^2_\nu)})$, and θ_3 is the coefficient corresponding to $(\bar{J}_T \otimes \bar{J}_N)$, which can be verified to be $(1 - \sqrt{c_1/c_2} - \sqrt{c_1\sigma^2_\nu/(T\sigma^2_\mu + c_1\sigma^2_\nu)} + \sqrt{c_1\sigma^2_\nu/(T\sigma^2_\mu + c_2\sigma^2_\nu)})$. If $\lambda = 0$, then $\theta_1 = \theta_3 = 0$ and $\theta_2 = 1 - \sigma_\nu/\sqrt{T\sigma^2_\mu + \sigma^2_\nu}$, which reduces y^*_{ti} to $(y_{ti} - \theta_2 \bar{y}_{.i})$. This is the familiar Fuller and Battese (1973) random effects transformation that allows us to obtain GLS as weighted least squares.

If $\sigma^2_\mu = 0$, i.e., the case of no random effects, the variance-covariance matrix in (7) reduces to

$$\Omega = \sigma^2_\nu (I_T \otimes (B'_N B_N)^{-1}), \tag{16}$$

which from (11) reduces to

$$\Omega = \sigma^2_\nu (I_T \otimes (c_1 E_N + c_2 \bar{J}_N)). \tag{17}$$

Hence

$$\Omega^{-1} = \frac{1}{c_1\sigma^2_\nu}(I_T \otimes E_N) + \frac{1}{c_2\sigma^2_\nu}(I_T \otimes \bar{J}_N) \tag{18}$$

and

$$\Omega^{-1/2} = \frac{1}{\sigma_\nu\sqrt{c_1}}(I_T \otimes E_N) + \frac{1}{\sigma_\nu\sqrt{c_2}}(I_T \otimes \bar{J}_N), \tag{19}$$

so that the typical element of $y^* = \sigma_\nu\sqrt{c_1}\Omega^{-1/2}y$ is $y^*_{ti} = (y_{ti} - \theta_1 \bar{y}_{t.})$ where $\theta_1 = (1 - \sqrt{c_1/c_2})$. Note that as $N \to \infty$, $\sqrt{c_1/c_2} \to (1 - \lambda)$, and $\theta_1 \to \lambda$.

A special case of this model is the *cross-section* spatial regression model with $(T = 1)$ and equal weight matrix given by (8). In this case, one can show that OLS is equivalent to GLS as long as there is a constant in the regression. To prove this, note that the model in (5) becomes

$$y_N = \alpha \iota_N + X_N \beta + u_N = Z_N \gamma + u_N, \tag{20}$$

where $Z_N = (\iota_N, X_N)$ and $\gamma = (\alpha, \beta)'$. Here y_N is a vector of observations on the dependent variable and ι_N is a vector of ones, both of dimension N, and X_N

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

is an $N \times K$ matrix of observations on the K explanatory variables. The disturbance vector is assumed to follow a spatial autoregressive process

$$u_N = \lambda W_N u_N + \nu_N, \tag{21}$$

where W_N is an $N \times N$ equal weighting matrix given in (8). In fact, one can prove that the Zyskind (1967) necessary and sufficient condition for OLS to be equivalent to GLS on (20) is satisfied. This calls for $P_Z\Sigma = \Sigma P_Z$, where $Z = Z_N = (\iota_N, X_N)$ is the matrix of regressors in (20) and $\Sigma = E(u_N u_N') = \sigma_\nu^2 \Omega$ is the variance-covariance matrix of the disturbances given in (21). It is straightforward to show that $\Omega = c_1 E_N + c_2 \bar{J}_N$ and

$$P_Z\Omega = \frac{(N-1)^2}{(N-1+\lambda)^2}(P_Z - \bar{J}_N) + \frac{1}{(1-\lambda)^2}\bar{J}_N \tag{22}$$

because $P_Z\iota_N = \iota_N$, $P_Z\bar{J}_N = \bar{J}_N$, and $P_Z E_N = P_Z - \bar{J}_N$. Similarly,

$$\Omega P_Z = \frac{(N-1)^2}{(N-1+\lambda)^2}(P_Z - \bar{J}_N) + \frac{1}{(1-\lambda)^2}\bar{J}_N. \tag{23}$$

Hence, $P_Z\Omega = \Omega P_Z$.

In fact, another necessary and sufficient condition for OLS to be equivalent to GLS, which relies on Ω^{-1}, is given by Milliken and Albohali (1984), and this condition calls for $Z'\Omega^{-1}(I_N - P_Z) = 0$. Here

$$\Omega^{-1} = \frac{(N-1+\lambda)^2}{(N-1)^2}E_N + (1-\lambda)^2\bar{J}_N, \tag{24}$$

and using the fact that $\bar{J}_N(I_N - P_Z) = 0$, we get

$$\begin{aligned}\Omega^{-1}(I_N - P_Z) &= \frac{(N-1+\lambda)^2}{(N-1)^2}E_N(I_N - P_Z) + (1-\lambda)^2\bar{J}_N(I_N - P_Z) \\ &= \frac{(N-1+\lambda)^2}{(N-1)^2}(I_N - P_Z).\end{aligned} \tag{25}$$

Hence

$$Z'\Omega^{-1}(I_N - P_Z) = \frac{(N-1+\lambda)^2}{(N-1)^2}Z'(I_N - P_Z) = 0 \tag{26}$$

because $Z'P_Z = Z'$. Note that

$$\begin{aligned}\mathrm{var}(\hat{\gamma}_{OLS}) &= \mathrm{var}(\hat{\gamma}_{GLS}) = \sigma_\nu^2(Z'\Omega^{-1}Z)^{-1} \\ &= \sigma_\nu^2\left[(1-\lambda)^2(Z'\bar{J}_N Z) + \frac{(N-1+\lambda)^2}{(N-1)^2}(Z'E_N Z)\right]^{-1}\end{aligned} \tag{27}$$

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

and this, in general, is *not* equal to the usual formula for $\text{var}(\hat{\gamma}_{OLS})$ computed by regression packages, i.e., $\sigma_\nu^2(Z'Z)^{-1}$, unless $\lambda = 0$, which is the case of no spatial correlation.

For the spatial *panel* regression with equal weights, this result is not necessarily true as long as $\sigma_\mu^2 > 0$. Two special cases where OLS on (5) is the same as GLS, i.e., $\hat{\gamma}_{OLS} = (Z'Z)^{-1}Z'y = \hat{\gamma}_{GLS} = (Z'\Omega^{-1}Z)^{-1}Z'\Omega^{-1}y$, are the following: (i) the trivial case where $\sigma_\mu^2 = 0$ and $\lambda = 0$ (in fact, when $\lambda = 0$, $c_1 = c_2 = 1$, $\theta_1 = 0$, and $y_{ti}^* = y_{ti}$ and GLS reduces to OLS); also, (ii) when the matrix of regressors X is invariant across time, i.e., when $X = \iota_T \otimes X_N$, with X_N being the $N \times k$ matrix of exogenous regressors that is invariant across time. To show this, we prove that the Zyskind (1967) necessary and sufficient condition for OLS to be equivalent to GLS on (5) is satisfied. This calls for $P_Z\Omega = \Omega P_Z$, where $Z = \iota_T \otimes Z_N = \iota_T \otimes (\iota_N, X_N)$ is the matrix of regressors in (5) and Ω is the variance-covariance matrix of the disturbances given in (17). It is straightforward to show that $P_Z = \bar{J}_T \otimes P_{Z_N}$ and that $P_{Z_N}\iota_N = \iota_N$, $P_{Z_N}\bar{J}_N = \bar{J}_N$, and $P_{Z_N}E_N = P_{Z_N} - \bar{J}_N$. Hence

$$\begin{aligned} P_Z\Omega &= \sigma_\nu^2(\bar{J}_T \otimes P_{Z_N})(I_T \otimes (c_1 E_N + c_2 \bar{J}_N)) \\ &= \sigma_\nu^2(\bar{J}_T \otimes (c_1(P_{Z_N} - \bar{J}_N) + c_2 \bar{J}_N)). \end{aligned} \tag{28}$$

Similarly,

$$\begin{aligned} \Omega P_Z &= \sigma_\nu^2(I_T \otimes (c_1 E_N + c_2 \bar{J}_N))(\bar{J}_T \otimes P_{Z_N}) \\ &= \sigma_\nu^2(\bar{J}_T \otimes (c_1(P_{Z_N} - \bar{J}_N) + c_2 \bar{J}_N)). \end{aligned} \tag{29}$$

Hence, $P_Z\Omega = \Omega P_Z$.

In fact, another necessary and sufficient condition for OLS to be equivalent to GLS, which relies on Ω^{-1}, is given by Milliken and Albohali (1984), and this condition calls for $Z'\Omega^{-1}(I_{TN} - P_Z) = 0$. Using (18), and the fact that $\bar{J}_N(I_N - P_{Z_N}) = 0$, we get

$$\begin{aligned} \Omega^{-1}(I_{TN} - P_Z) &= \left(\frac{1}{c_1\sigma_\nu^2}(I_T \otimes E_N) + \frac{1}{c_2\sigma_\nu^2}(I_T \otimes \bar{J}_N)\right)(I_{TN} - P_Z) \\ &= \left(\frac{1}{c_1\sigma_\nu^2}(I_T \otimes E_N) + \frac{1}{c_2\sigma_\nu^2}(I_T \otimes \bar{J}_N)\right) \\ &\quad - \left(\frac{1}{c_1\sigma_\nu^2}(\bar{J}_T \otimes (P_{Z_N} - \bar{J}_N)) + \frac{1}{c_2\sigma_\nu^2}(\bar{J}_T \otimes \bar{J}_N)\right) \\ &= \frac{1}{c_1\sigma_\nu^2}((I_T \otimes E_N) - (\bar{J}_T \otimes (P_{Z_N} - \bar{J}_N))) \\ &\quad + \frac{1}{c_2\sigma_\nu^2}(E_T \otimes \bar{J}_N). \end{aligned} \tag{30}$$

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

Hence

$$Z'\Omega^{-1}(I_{TN} - P_Z) = \frac{1}{c_1\sigma_\nu^2}\{(\iota_T' \otimes Z_N')((I_T \otimes E_N) - (\bar{J}_T \otimes (P_{Z_N} - \bar{J}_N)))\}$$

$$= \frac{1}{c_1\sigma_\nu^2}\{(\iota_T' \otimes Z_N'E_N) - (\iota_T' \otimes Z_N'(P_{Z_N} - \bar{J}_N))\}$$

$$= \frac{1}{c_1\sigma_\nu^2}\{(\iota_T' \otimes Z_N'E_N) - (\iota_T' \otimes (Z_N' - Z_N'\bar{J}_N))\} = 0 \qquad \textbf{(31)}$$

because $\iota_T'\bar{J}_T = \iota_T'$ and $Z_N'P_{Z_N} = Z_N'$.

An anonymous referee pointed out that an alternative derivation of these results can be obtained using Lemma 2.1 of Magnus (1982). This lemma was derived in the context of a multivariate error component model, but when applied to the model considered in this paper, one common theme is the following: if $W_N = \alpha_1 A_N + \alpha_2(I_N - A_N)$ with $\alpha_1 \neq \alpha_2$, and A_N is a symmetric idempotent matrix, then all powers of W_N (also negative powers), its eigenvalues, and determinant can all be easily calculated. Here, we have $I_N - \lambda W_N = \beta_1 A_N + \beta_2(I_N - A_N)$, with $\beta_1 = 1 - \lambda\alpha_1$ and $\beta_2 = 1 - \lambda\alpha_2$. Assuming that $\lambda\alpha_1 \neq 1$ and $\lambda\alpha_2 \neq 1$, we get $(I_N - \lambda W_N)^{-1} = (1/\beta_1)A_N + (1/\beta_2)(I_N - A_N)$. Hence $\Omega = (I_N - \lambda W_N)^{-2} = (1/\beta_1^2)A_N + (1/\beta_2^2)(I_N - A_N)$. Applying the Milliken and Albohali (1984) necessary and sufficient condition that calls for $Z'\Omega^{-1}(I_N - P_Z) = 0$, we get $Z'\Omega^{-1}(I_N - P_Z) = (\beta_1^2 - \beta_2^2)Z'A_N(I_N - P_Z) = 0$, which is $\Leftrightarrow Z'A_N(I_N - P_Z) = 0$. In the special case where $A_N = \bar{J}_N$, we see that OLS is equivalent to GLS if and only if $Z'\iota_N = 0$ or $(I_N - P_Z)\iota_N = 0$, i.e., if and only if the regressors are measured in deviations from the mean or the regression contains a constant.

3. EXTENSIONS

So far, we have considered the *spatial autoregressive* (SAR) specification for the disturbances. An alternative specification is the *spatial moving average* (SMA) specification. Anselin (2003) classifies the spatial covariance structure induced by SAR as *global* and that by SMA as *local*. In this case, the model given by (1) and (2) is the same, but the disturbances in (3) become

$$\epsilon_t = \lambda W_N \nu_t + \nu_t = (I_N + \lambda W_N)\nu_t \qquad \textbf{(32)}$$

In vector form, the panel disturbances in (6) become

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes (I_N + \lambda W_N))\nu \qquad \textbf{(33)}$$

with variance covariance matrix

$$\Omega = \sigma_\mu^2(J_T \otimes I_N) + \sigma_\nu^2(I_T \otimes (I_N + \lambda W_N)^2) \qquad \textbf{(34)}$$

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

because W_N given by (9) is symmetric. In fact,

$$(I_N + \lambda W_N)^2 = \frac{(N-1-\lambda)^2}{(N-1)^2} E_N + (1+\lambda)^2 \bar{J}_N = d_1 E_N + d_2 \bar{J}_N, \tag{35}$$

where $d_1 = (N-1-\lambda)^2/(N-1)^2$ and $d_2 = (1+\lambda)^2$.

Replacing J_T and I_T by their equivalent terms $T\bar{J}_T$ and $(E_T + \bar{J}_T)$ in (34), and collecting like terms, we obtain

$$\Omega = (T\sigma_\mu^2 + d_1\sigma_\nu^2)(\bar{J}_T \otimes E_N) + (T\sigma_\mu^2 + d_2\sigma_\nu^2)(\bar{J}_T \otimes \bar{J}_N) + d_1\sigma_\nu^2(E_T \otimes E_N) + d_2\sigma_\nu^2(E_T \otimes \bar{J}_N). \tag{36}$$

Hence, we get the same results for the SMA specification as for the SAR specification when the weight matrix is equal. The only difference between (36) and (13) is d_1 and d_2 rather than c_1 and c_2. Similarly, Ω^{-1} and $\Omega^{-1/2}$ are the same as (14) and (15) with d_1 and d_2 replacing c_1 and c_2. The same holds true for the Fuller and Battese type transformation described following (15). Note that if $\lambda = 0$, so that there is no spatial autocorrelation, then $d_1 = d_2 = 1$ and Ω in (36) reduces to the familiar variance-covariance matrix of the random effects panel data model. If $\sigma_\mu^2 = 0$, i.e., the case of no random effects, the variance-covariance matrix in (34) reduces to

$$\Omega = \sigma_\nu^2(I_T \otimes (d_1 E_N + d_2 \bar{J}_N)). \tag{37}$$

Hence, Ω^{-1} and $\Omega^{-1/2}$ are the same as in (18) and (19) with d_1 and d_2 replacing c_1 and c_2. The Fuller and Battese transformation given following (19) is now $y^* = \sigma_\nu\sqrt{d_1}\Omega^{-1/2}y$ with typical element $y_{ti}^* = (y_{ti} - \theta_1\bar{y}_{t.})$ where $\theta_1 = (1 - \sqrt{d_1/d_2})$. Note that as $N \to \infty$, $\sqrt{d_1/d_2} \to 1/(1+\lambda)$ and $\theta_1 \to \lambda/(1+\lambda)$.

Kelejian and Robinson (1995) considered an alternative *spatial error components* (SEC) specification that differs from the SAR and SMA specifications. In this case, the model given by (1) and (2) is the same, but the disturbances in (3) become

$$\epsilon_t = \lambda W_N \psi_t + \nu_t, \tag{38}$$

where ψ_t is an $(N \times 1)$ vector of spillover error components. The two component vectors ψ_t and ν_t are assumed to consist of independent and identically distributed (i.i.d.) terms with respective variances σ_ψ^2 and σ_ν^2 and are uncorrelated. In vector form, the panel disturbances in (6) become

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes \lambda W_N)\psi + \nu \tag{39}$$

with variance-covariance matrix

$$\Omega = \sigma_\mu^2(J_T \otimes I_N) + (I_T \otimes (\sigma_\nu^2 I_N + \lambda^2\sigma_\psi^2 W_N^2)) \tag{40}$$

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

because W_N given by (9) is symmetric. In fact,

$$(\sigma_\nu^2 I_N + \lambda^2\sigma_\psi^2 W_N^2) = \left(\sigma_\nu^2 + \frac{\lambda^2\sigma_\psi^2}{(N-1)^2}\right)E_N + (\sigma_\nu^2 + \lambda^2\sigma_\psi^2)\bar{J}_N = b_1 E_N + b_2 \bar{J}_N, \tag{41}$$

where $b_1 = (\sigma_\nu^2 + (\lambda^2\sigma_\psi^2/(N-1)^2))$ and $b_2 = (\sigma_\nu^2 + \lambda^2\sigma_\psi^2)$. The rest of the derivations are the same as before with b_1 and b_2 replacing c_1 and c_2. Note that as $N \to \infty$, $\sqrt{b_1/b_2} \to \sigma_\nu/\sqrt{\sigma_\nu^2 + \lambda^2\sigma_\psi^2}$ and $\theta_1 \to 1 - \sigma_\nu/\sqrt{\sigma_\nu^2 + \lambda^2\sigma_\psi^2}$.

For the cross-section ($T = 1$) spatial regression model with SMA or SEC disturbances and an equal weight matrix, one can easily show that OLS is equivalent to GLS as long as there is a constant in the model; the proof is left to the reader. For the spatial panel regression with equal weights, this result is not necessarily true as long as $\sigma_\mu^2 > 0$. Two cases where OLS is the same as GLS are once again (i) the trivial case where $\sigma_\mu^2 = 0$ and $\lambda = 0$ or (ii) when the matrix of regressors X is invariant across time. The proofs for the SMA and SEC cases are the same as that for the SAR case and are left for the reader. The key for these results is the fact that (11) for SAR, (35) for SMA, and (41) for SEC are all linear combinations of E_N and $\bar{J}_N$. These matrices are idempotent, are orthogonal to each other, and sum to I_N. Alternatively, one can apply Lemma 2.1 of Magnus (1982) to obtain the same results.

Next, we consider an alternative *panel* data model with spatially correlated error components suggested by Kapoor et al. (2005). The regression model is the same as (1), but the SEC structure given by (2) and (3) becomes

$$u = \lambda(I_T \otimes W_N)u + \epsilon \tag{42}$$

with

$$\epsilon = (\iota_T \otimes I_N)\mu + \nu, \tag{43}$$

where μ and ν are the same as before, i.e., $iid(0, \sigma_\mu^2)$ and $iid(0, \sigma_\nu^2)$ independent of each other and among themselves. Performing the spatial Cochrane–Orcutt type transformation on (5) one gets

$$y^*(\lambda) = \alpha(1-\lambda)\iota_{NT} + X^*(\lambda)\beta + u^*(\lambda) = Z^*(\lambda)\gamma + u^*(\lambda) \tag{44}$$

with

$$u^*(\lambda) = (I_T \otimes (I_N - \lambda W_N))u = \epsilon \tag{45}$$

and $y^*(\lambda)$, $X^*(\lambda)$, and $Z^*(\lambda)$ defined similarly. The variance-covariance matrix $\Omega_\epsilon = E(\epsilon\epsilon') = (T\sigma_\mu^2 + \sigma_\nu^2)(\bar{J}_T \otimes I_N) + \sigma_\nu^2(E_T \otimes I_N)$ is the usual panel data error components random effects variance-covariance matrix. GLS on (44) can be obtained as OLS on the Fuller–Battese transformed equation (44), i.e., after premultiplying it by $\sigma_\nu \Omega_\epsilon^{-1/2} = (E_T \otimes I_N) + (\sigma_\nu/\sigma_1)(\bar{J}_T \otimes I_N)$, where $\sigma_1 = \sqrt{T\sigma_\mu^2 + \sigma_\nu^2}$. In this case,

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

$$y^{**}(\lambda) = \sigma_\nu \Omega_\epsilon^{-1/2} y^*(\lambda) = \left((E_T \otimes I_N) + \frac{\sigma_\nu}{\sigma_1} (\bar{J}_T \otimes I_N) \right) (I_T \otimes (I_N - \lambda W_N)) y. \tag{46}$$

For the equal weight matrix given in (8), $(I_N - \lambda W_N)$ is given following (9) and can be rewritten as $(I_N - \lambda W_N) = a_1 E_N + a_2 \bar{J}_N$, where $a_1 = (N - 1 + \lambda)/(N - 1)$ and $a_2 = (1 - \lambda)$. Substituting this expression in (46) yields

$$\begin{aligned} y^{**}(\lambda) &= \left[(E_T \otimes (a_1 E_N + a_2 \bar{J}_N)) + \frac{\sigma_\nu}{\sigma_1} (\bar{J}_T \otimes (a_1 E_N + a_2 \bar{J}_N)) \right] y \\ &= a_1 (E_T \otimes E_N) y + a_2 (E_T \otimes \bar{J}_N) y \\ &\quad + \frac{a_1 \sigma_\nu}{\sigma_1} (\bar{J}_T \otimes E_N) y + \frac{a_2 \sigma_\nu}{\sigma_1} (\bar{J}_T \otimes \bar{J}_N) y. \end{aligned} \tag{47}$$

Once again, we see that for the Kapoor et al. (2005) specification, with an equal weight matrix, GLS on (44) can be written as a weighted combination similar to that following (15), i.e., $y_{ti}^{**} = a_1 (y_{ti} - \theta_1 \bar{y}_{t.} - \theta_2 \bar{y}_{.i} + \theta_3 \bar{y}_{..})$ where $\bar{y}_{t.}$ denotes the sample average over regions, $\bar{y}_{.i}$ denotes the sample average over time, and $\bar{y}_{..}$ denotes the average over the entire sample. The θ's can be easily obtained from the corresponding $(a_1, a_2, \sigma_\mu, \sigma_\nu)$ parameters. For example, θ_1 is the coefficient corresponding to $(I_T \otimes \bar{J}_N)$, which can be verified to be $(1 - (a_2/a_1))$, θ_2 is the coefficient corresponding to $(\bar{J}_T \otimes I_N)$, which can be verified to be $(1 - (\sigma_\nu/\sigma_1))$, and θ_3 is the coefficient corresponding to $(\bar{J}_T \otimes \bar{J}_N)$, which can be verified to be $(1 - (a_2/a_1) - (\sigma_\nu/\sigma_1) + (a_2 \sigma_\nu / a_1 \sigma_1))$. If $\lambda = 0$, $a_1 = a_2 = 1$, then $\theta_1 = \theta_3 = 0$ and $\theta_2 = (1 - (\sigma_\nu/\sigma_1))$, which reduces y_{ti}^{**} to $(y_{ti} - \theta_2 \bar{y}_{.i})$. This is the familiar Fuller and Battese (1973) random effects transformation.

If $\sigma_\mu^2 = 0$, i.e., the case of no random effects, then $\sigma_\nu = \sigma_1$, $\theta_2 = \theta_3 = 0$, and $\theta_1 = (1 - (a_2/a_1))$, which reduces y_{ti}^{**} to $a_1 (y_{ti} - \theta_1 \bar{y}_{t.})$. In fact, $y^{**}(\lambda)$ from (47) reduces to

$$\begin{aligned} y^{**}(\lambda) &= [I_T \otimes (a_1 E_N + a_2 \bar{J}_N)] y = a_1 (I_T \otimes E_N) y + a_2 (I_T \otimes \bar{J}_N) y \\ &= a_1 y + (a_2 - a_1)(I_T \otimes \bar{J}_N) y. \end{aligned} \tag{48}$$

Note that $a_1 \theta_1 = (a_1 - a_2)$. Therefore, y_{ti}^{**} is a weighted combination of y_{ti} and $\bar{y}_{t.}$. For this model, OLS is the same as GLS for the trivial case where $\sigma_\mu^2 = 0$ and $\lambda = 0$. The more interesting case where OLS on (44) is equivalent to GLS is the case where the matrix of regressors X is invariant across time. The proof is along the same lines as before and is left for the reader. In fact, one can easily show that when $X = \iota_T \otimes X_N$, with X_N being the $N \times k$ matrix of exogenous regressors that is invariant across time, both the Zyskind (1967) and the Milliken and Albohali (1984) necessary and sufficient conditions are satisfied for this model.

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press

4. CONCLUSION

For spatial panel models with a weight matrix that is row normalized and has equal weights within a block but otherwise uncorrelated across blocks, we showed that GLS can be obtained as a simple weighted least squares transformation involving the means of the observations over time and over units. This is similar to the Fuller and Battese (1973) transformation for the random effects model in panel data with no spatial correlation. In addition, we showed that OLS is equivalent to GLS if there are no random or spatial effects or if the regressors vary across units but are invariant over time. We also showed that these results for the equal weight matrix hold whether we use the SAR specification, the SMA specification, the SEC specification, or the alternative panel data regression model with spatially correlated error components suggested by Kapoor et al. (2005).

REFERENCES

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

Anselin, L. (2003) Spatial externalities, spatial multipliers and spatial econometrics. *International Regional Science Review* 26, 153–166.

Baltagi, B.H. (2005) *Econometric Analysis of Panel Data*. Wiley.

Baltagi, B.H., S.H. Song, & W. Koh (2003) Testing panel data regression models with spatial error correlation. *Journal of Econometrics* 117, 123–150.

Case, A.C. (1991) Spatial patterns in household demand. *Econometrica* 59, 953–965.

Fuller, W.A. & G.E. Battese (1973) Transformations for estimation of linear models with nested error structure. *Journal of the American Statistical Association* 68, 626–632.

Holtz-Eakin, D. (1994) Public-sector capital and the productivity puzzle. *Review of Economics and Statistics* 76, 12–21.

Kapoor, M., H.H. Kelejian, & I.R. Prucha (2005) Panel data models with spatially correlated error components. *Journal of Econometrics*, forthcoming.

Kelejian, H.H. & I.R. Prucha (2002) 2SLS and OLS in a spatial autoregressive model with equal spatial weights. *Regional Science and Urban Economics* 32, 691–707.

Kelejian, H.H., I.R. Prucha, & Y. Yuzefovich (2005) Estimation problems with spatial weighting matrices which have blocks of equal elements. *Journal of Regional Science,* forthcoming.

Kelejian, H.H. & D.P. Robinson (1995) Spatial correlation: A suggested alternative to the autoregressive model. In L. Anselin & R.J. Florax (eds.), *New Directions in Spatial Econometrics*, pp. 75–95. Springer-Verlag.

Lee, L.F. (2002) Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models. *Econometric Theory* 18, 252–277.

Magnus, J.R. (1982) Multivariate error components analysis of linear and nonlinear regression models by maximum likelihood. *Journal of Econometrics* 19, 239–285.

Milliken, G.A. & M. Albohali (1984) On necessary and sufficient conditions for ordinary least squares estimators to be best linear unbiased estimators. *American Statistician* 38, 298–299.

Wansbeek, T.J. & A. Kapteyn (1982) A simple way to obtain the spectral decomposition of variance components models for balanced data. *Communications in Statistics* A11, 2105–2112.

Zyskind, G. (1967) On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *Annals of Mathematical Statistics* 36, 1092–1109.

https://doi.org/10.1017/S0266466606060464 Published online by Cambridge University Press