Regional Science and Urban Economics 32 (2002) 691–707
regional SCIENCE & urban ECONOMICS
www.elsevier.com/locate/econbase

2SLS and OLS in a spatial autoregressive model with equal spatial weights

Harry H. Kelejian*, Ingmar R. Prucha

Department of Economics, University of Maryland, College Park, MD 20742, USA

Received 14 November 2000; received in revised form 4 December 2001; accepted 4 January 2002

Abstract

The paper considers a Cliff–Ord type spatial model with a spatially lagged dependent variable and a row normalized weighting matrix with equal weights. We show that the 2SLS and OLS estimators are inconsistent unless panel data are available. The weighting matrix in question is one which would naturally be considered if all units are neighbors to each other, and there is no other reasonable or observable measure of distance between them.

Keywords: Spatial autoregressive models; Row normalized and equal spatial weights; Ordinary least squares; Two stage least squares; Panel data

JEL classification: C0; C2

1. Introduction

We consider estimation issues in linear Cliff–Ord type spatial models which contain a spatially lagged dependent variable.[1] We give both a 'negative' result that serves as a warning for a potential pitfall, as well as a 'positive result'. More

*Corresponding author. Tel.: +1-301-405-3492; fax: +1-301-405-3542.

E-mail address: kelejian@econ.umd.edu (H.H. Kelejian).

[1] Basic references on spatial models are Anselin (1988), Cliff and Ord (1973, 1981), and Cressie (1993). A selection of more recent studies which relate to certain theoretical and empirical issues are Case (1991), Boarnet (1994), Anselin et al. (1996), Bollinger and Ihlanfeldt (1997), Pace and Barry (1977), Conley (1999), Lee (1999a,b,c), and Kelejian and Prucha (1998, 1999, 2001).

0166-0462/02/$ – see front matter
PII: S0166-0462(02)00003-0

specifically, we first show that under typical conditions the two-stage least squares (2SLS) and the ordinary least squares (OLS) estimators are inconsistent if the weighting matrix is row normalized and has equal weights, and data are only available for a single cross section. We then show that this problem typically does not occur if panel data are available. Ironically, if two or more panels of data are available, both the 2SLS and OLS estimator are not only consistent but are also efficient within the class of instrumental variable estimators.[2] These results are important to note because the weighting matrix considered is one which would naturally suggest itself if all units are neighbors to each other and there is no other natural or observable measure of distance. Cases which may be consistent with this are ones in which all cross sectional units interact in a confined space. Such a matrix was considered by Splitstoser (1999) in a study of spatial interdependence involving the ideology of legislators. It was also considered by Case (1992) in a panel data study of the adoption of new technologies by farmers, and by Lee (1999b) in a study of the properties of least squares estimators in linear spatial models.[3]

The paper is organized as follows. In Section 2 we specify the considered Cliff–Ord type model for a single cross section, and demonstrate the inconsistency of 2SLS and OLS within this setting. Section 3 extends the framework to panel data and demonstrates that in contrast to the previous case, the model parameters can now be estimated consistently by 2SLS and OLS. We also derive the limiting distribution of the estimators. All proofs are relegated to Appendix A.

The following notation and conventions are useful: let v_N and P_N be, respectively, a vector and matrix. Then we denote the ith element of v_N as $v_{i,N}$ and the i,jth element of P_N as $p_{ij,N}$. The same convention holds for vectors and matrices which do not involve the index N, in which case the index N is suppressed. We will say that the elements of the sequence of matrices P_N are uniformly bounded in absolute value if

$$|p_{ij,N}| \leq C < \infty$$

for all $1 \leq i, \ j \leq N$; $N \geq 1$, where the constant C does not depend on any of the indices.

2. 2SLS and OLS in the case of a single cross section

2.1. The model

Consider the following cross sectional spatial model ($N \geq 1$)

[2] The OLS estimator may, of course, be viewed as a special case of an instrumental variable estimator.

[3] Other relevant cases may relate to student performance in a school, certain interactions of workers in a firm, etc.

$$\begin{aligned} y_N &= \alpha e_N + X_N\beta + \lambda W_N y_N + \varepsilon_N \\ &= Z_N\gamma + \varepsilon_N \end{aligned} \tag{1}$$

where $Z_N = (e_N, X_N, W_N y_N)$ and $\gamma = (\alpha, \beta', \lambda)'$. Here y_N is the $N \times 1$ vector of observations on the dependent variable, e_N is the $N \times 1$ vector of unit elements, X_N is an $N \times k$ matrix of observations on k exogenous variables, W_N is an $N \times N$ weighting matrix, ε_N is the disturbance vector, α and λ are scalar parameters, and β is a $k \times 1$ vector of parameters. Our analysis is conditional on the realized values of the exogenous variables and so the matrices X_N will be taken as matrices of constants.

Our assumptions for (1) are the standard ones, except for the specification of the weighting matrix. Let $D_N = (e_N, X_N)$. Then, we assume

Assumption 1. $|\lambda| < 1$. Also, for all $1 \le i \le N$, $N \ge 1$ the error terms $\varepsilon_{i,N}$ are identically distributed with mean zero and finite variance σ^2. In addition, for each N the error terms $\varepsilon_{1,N}, \ldots, \varepsilon_{N,N}$ are independently distributed.

Assumption 2. The elements of the sequence of matrices X_N are uniformly bounded in absolute value, and $Q_{DD} = \lim_{N\to\infty} N^{-1} D_N' D_N$ is finite and nonsingular.

It proves convenient to decompose Q_{DD} as

$$Q_{DD} = \begin{bmatrix} 1 & \mu_x' \\ \mu_x & Q_{XX} \end{bmatrix}.$$

Clearly μ_x is a $k \times 1$ vector, and Q_{XX} is a $k \times k$ matrix. Let $\bar{x}_N = N^{-1} X_N' e_N$ be the vector of sample means of the exogenous regressors. Then for future reference note that given the above assumption $\bar{x}_N \to \mu_x$ and $N^{-1} X_N' X_N \to Q_{XX}$.

Assumption 3. $W_N = (1/(N-1))\ J_N - (1/(N-1))\ I_N$, $N > 1$, where $J_N = e_N e_N'$. $W_1 = 0$.

This assumption implies that

$$w_{ij,N} = \begin{cases} \dfrac{1}{N-1} & \text{for} \quad i \neq j \\ 0 & \text{for} \quad i = j \end{cases}$$

and thus, among other things, the model in (1) relates each element of y_N to the average of the other $N-1$ elements.

Remark. *It seems of interest to briefly compare the above spatial autoregressive model with the widely used class of mixed models that combines fixed and random effects. Following, e.g. Harville* (1988), *a mixed model is defined as*

$$y_N = F_N\delta + R_N s_N + \varepsilon_N$$

where y_N denotes the vector of observations on the dependent variable, F_N and R_N are known nonstochastic matrices (e.g. observations on exogenous refressors), δ is a vector of unknown parameters, s_N is a vector of unobserved random effects with zero mean, ε_N is the unobserved disturbance vector, and s_N and ε_N are uncorrelated. In (1) $\lambda W_N y_N$ can be expressed as $W_N(\lambda y_N)$. Therefore, if we define $F_N = (e_N, X_N)$, $\delta = (\alpha, \beta')'$, $R_N = W_N$ and $s_N = \lambda y_N$ it may seem, at first glance, that we can express the spatial autoregressive model (1) in the above form of a mixed model. However, upon closer inspection we see that the two models are substantially different. Among other things, in (1) the 'random effects' $s_N = \lambda y_N$ are observed (up to a common scalar), have nonzero mean, and clearly $s_N = \lambda y_N$ and ε_N are correlated. This is reflective of the fact that the spatial autoregressive model represents a simultaneous system, while the above mixed model does not. The subsequent discussion pertains to the spatial autoregressive model as defined before this paragraph.

2.2. Some useful results

For future reference we note that, as is easily verified,

$$(I - \lambda W_N)^{-1} = \delta_{1N} J_N + \delta_{2N} I_N \tag{2}$$

where

$$\delta_{1N} = \frac{\lambda}{(N-1+\lambda)(1-\lambda)} \quad \text{and} \quad \delta_{2N} = \frac{N-1}{N-1+\lambda}. \tag{3}$$

The result in (2) can be used to obtain the reduced form for the spatial lag of the dependent variable in (1). Specifically, (1) implies that $W_N y_N = W_N(I - \lambda W_N)^{-1}[\alpha e_N + X_N\beta + \varepsilon_N]$ and so, in light of (2),

$$W_N y_N = c_N e_N + \phi_{2N} X_N \beta + \phi_{2N} \varepsilon_N \tag{4}$$

where

$$c_N = N\phi_{1N}[\alpha + \bar{x}'_N\beta + \bar{\varepsilon}_N] + \alpha\phi_{2N},$$
$$\phi_{1N} = \delta_{1N} + \frac{\delta_{2N}}{N-1}, \quad \phi_{2N} = -\frac{\delta_{2N}}{N-1},$$

and $\bar{\varepsilon}_N = N^{-1}e'_N\varepsilon_N$. From the expressions in (3) we see that

$$\phi_{1N} \to 0, \quad \phi_{2N} \to 0 \text{ and } N\phi_{1N} \to 1/(1-\lambda) \tag{5}$$

as $N \to \infty$.

Since $\bar{\varepsilon}_N$ has zero mean and variance σ^2/N, it follows from Chebyshev's inequality that $p\lim_{N\to\infty} \bar{\varepsilon}_N = 0$. Therefore, again for future reference, we note that

$$p\lim_{N\to\infty} c_N = \frac{\alpha + \mu'_x\beta}{1-\lambda}. \tag{6}$$

The results in (4)–(6) suggest that there may be a basic estimation problem in that, asymptotically, the spatial lag $W_N y_N$ is proportional to the unit vector, and thus collinear with the intercept. In the following sections we discuss this problem and its effect on the 2SLS and OLS estimator in more detail.

2.3. 2SLS estimation

Since $W_N y_N$ is correlated with the disturbance term in (1), see, e.g., Anselin (1988), and Kelejian and Prucha (1998), one suggested method of estimation for γ is 2SLS. Let $H_N = (e_N, X_N, G_N)$ be the matrix of instruments, where G_N is an $N \times r$ matrix of nonstochastic variables, $r \geq 1$. We maintain the following assumption concerning those instruments.

Assumption 4. The elements of the sequence of matrices G_N are uniformly bounded in absolute value, and $Q_{HH} = \lim_{N\to\infty} N^{-1} H_N' H_N$ is finite and nonsingular.

One suggestion for the columns of G_N given by Kelejian and Prucha (1998) and by Rey and Boarnet (1998) are the spatial lags of X_N of various orders. We note, however, that the subsequent discussion does not postulate any particular structure for G_N apart from what is maintained in Assumption 4.

Let $P_{H_N} = H_N (H_N' H_N)^{-1} H_N$ and note that since H_N contains e_N and X_N, $P_{H_N} e_N = e_N$ and $P_{H_N} X_N = X_N$. Let

$$\hat{Z}_N = P_{H_N} Z_N = (e_N, X_N, P_{H_N} W_N y_N). \tag{7}$$

The 2SLS estimator of γ is then given by

$$\hat{\gamma}_{2\mathrm{SLS},N} = (\hat{Z}_N' \hat{Z}_N)^{-1} \hat{Z}_N' y_N. \tag{8}$$

The following lemma, whose proof is given in Appendix A, establishes that the 2SLS estimator is inconsistent under the assumptions maintained above.

Lemma 1. *Given the model in* (1), *Assumptions* 1, 2, 3 *and* 4,

$$p \lim_{N\to\infty} \hat{\gamma}_{2\mathrm{SLS},N} \neq \gamma.$$

On an intuitive level the above lemma can be motivated from the following observations: results given in Amemiya (1985, pp. 245–255) imply that the optimal instrument matrix is $E(Z_N)$. In terms of our model, however, $E[\bar{\varepsilon}_N] = 0$ and so the results in (4) imply

$$E(Z_N) = [e_N, X_N, E(c_N) e_N + \phi_{2N} X_N \beta] \tag{9}$$

where

$$E(c_N) = \alpha N \phi_{1N} + N \phi_{1N} \bar{x}_N' \beta + \alpha \phi_{2N}.$$

Since $E(c_N)$, ϕ_{1N}, and ϕ_{2N} are scalars, the last column of $E(Z_N)$ is clearly multicollinear with the first $k+1$, and so $E(Z_N)$ does not have full column rank.

We also note that the above lemma does not contradict 'positive' results given in Kelejian and Prucha (1998) concerning the consistency and asymptotic normality of the 2SLS estimator in linear spatial models containing a spatially lagged dependent variable. For instance, Eqs. (4)–(6) imply that

$$Q_{HZ} = p \lim_{N\to\infty} N^{-1} H_N' Z_N = p \lim_{N\to\infty} N^{-1} H_N' \left[e_N, X_N, \left(\frac{\alpha + \mu_x' \beta}{1-\lambda} \right) e_N \right],$$

which shows that Q_{HZ} has less than full column rank. This violates Assumption 7 maintained in Kelejian and Prucha (1998). We note further that as a consequence

$$\begin{aligned} p \lim_{N\to\infty} N^{-1}(\hat{Z}_N' \hat{Z}_N) &= p \lim_{N\to\infty} [N^{-1} Z_N' H_N][N^{-1}(H_N' H_N)]^{-1}[N^{-1} H_N' Z_N] \\ &= Q_{HZ}' Q_{HH}^{-1} Q_{HZ} \end{aligned}$$

is singular.

2.4. OLS estimation

Lee (1999b) has shown that under certain conditions the OLS estimator of the parameters of a linear spatial model containing a spatially lagged dependent variable is consistent and asymptotically normal. One of the conditions assumed in Lee's model implies that each element of the spatial weights matrix limits to zero as the sample size $N \to \infty$. Since that condition is satisfied by our model, one might also consider the least squares estimator of γ in (1), i.e. $\hat{\gamma}_{\mathrm{OLS},N} = (Z_N' Z_N)^{-1} Z_N' y_N$. However, as seen from the subsequent lemma, under the maintained assumptions the OLS estimator is also inconsistent.

Lemma 2. *Given the model in* (1), *and Assumptions* 1, 2, 3,

$$p \lim_{N\to\infty} \hat{\gamma}_{\mathrm{OLS},N} \neq \gamma.$$

We note that the above lemma does not contradict the results given in Lee (1999b). While at first glance the spatial weights matrix defined in Assumption 3 seems to satisfy Lee's assumptions, one of Lee's conditions for the consistency of the OLS estimator is that $Q_{ZZ} = \lim_{N\to\infty} N^{-1} E(Z_N)' E(Z_N)$ is a finite nonsingular matrix. Under the assumptions maintained here this condition does not hold. Using (4)–(6) and (9) it is readily seen that

$$Q_{ZZ} = \begin{bmatrix} 1 & \mu_x' & c_* \\ \mu_x & Q_{XX} & c_* \mu_x \\ c_* & c_* \mu_x' & c_*^2 \end{bmatrix}$$

where $c_* = (\alpha + \mu_x'\beta)/(1-\lambda)$, and so Q_{ZZ} is singular since its first and third columns are proportional.

2.5. Limiting log-likelihood function

As was observed above, the root cause for the inconsistency of the 2SLS and OLS estimator of λ is that asymptotically $W_N y_N$ is proportional to the unit vector, and thus collinear with the intercept. This suggests that the problem encountered with the 2SLS and OLS estimator is not specific to those estimators. To shed more light on the issue it seems of interest to investigate the limiting log-likelihood function, properly normalized by the sample size.[4] Under normality the normalized log-likelihood function at parameter values $\underline{\alpha}$, $\underline{\beta}$, $\underline{\lambda}$, $\underline{\sigma}^2$ is given by

$$L_N(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2) = -\frac{1}{2}\ln(2\pi) + L_{1N}(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2) + L_{2N}(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2)$$

where

$$\begin{aligned} L_{1N}(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2) &= \frac{1}{N}\ln\|I_N - \underline{\lambda}W_N\| \\ &= \frac{N-1}{N}\ln\left(1 + \frac{\underline{\lambda}}{N-1}\right) + \frac{1}{N}\ln(1-\underline{\lambda}), \\ L_{2N}(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2) &= -\frac{1}{2}\ln(\underline{\sigma}^2) - \frac{1}{2\underline{\sigma}^2}\frac{1}{N}[y_N - Z_N\underline{\gamma}]'[y_N - Z_N\underline{\gamma}], \end{aligned}$$

where $\underline{\gamma} = (\underline{\alpha}, \underline{\beta}', \underline{\lambda})'$. The term $L_{1N}(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2)$ represents the log of the absolute value of the determinant of the Jacobian normalized by the sample size. We thank one of the referees for supplying us with the second expression for $L_{1N}(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2)$, which makes the maximum likelihood estimator easily computable. From this expression it is also readily seen that

$$\lim_{N\to\infty} L_{1N}(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2) = 0.$$

Thus this term drops out from the normalized log-likelihood function as the sample size tends to infinity. The term $L_{2N}(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2)$ is in essence the objective function of the OLS estimator. Observe that in light of (1) and the results in (4)–(6) we have

[4] This normalization ensures that the terms involved converge to finite limits; compare, e.g., Gallant (1987) and Pötscher and Prucha (1997) who provide a general analysis of the limiting behavior of extremum estimators.

$$
\begin{aligned}
& p\lim_{N\to\infty} \frac{1}{N}[y_N - Z_N\underline{\gamma}]'[y_N - Z_N\underline{\gamma}] \\
& \quad = p\lim_{N\to\infty} \frac{1}{N}[\varepsilon_N - Z_N(\underline{\gamma} - \gamma)]'[\varepsilon_N - Z_N(\underline{\gamma} - \gamma)] \\
& \quad = \sigma^2 + (\underline{\gamma} - \gamma)' Q_{ZZ}(\underline{\gamma} - \gamma).
\end{aligned}
$$

Thus the limiting normalized log-likelihood function is given by

$$
\begin{aligned}
p\lim_{N\to\infty} L_N(\underline{\alpha}, \underline{\beta}, \underline{\lambda}, \underline{\sigma}^2) = & -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\underline{\sigma}^2) \\
& - \frac{1}{2\underline{\sigma}^2}\left\{\sigma^2 + (\underline{\gamma} - \gamma)' Q_{ZZ}(\underline{\gamma} - \gamma)\right\}.
\end{aligned}
$$

As remarked above, the matrix Q_{ZZ} is singular, and thus the limiting normalized log-likelihood function does not have a unique maximum at the true parameter values α, β, λ, σ^2. In fact, as is readily seen, the function is maximized at any parameter value $\underline{\alpha}$, $\underline{\beta}$, $\underline{\lambda}$, $\underline{\sigma}^2$ with $\underline{\alpha} - \alpha + c_*(\underline{\lambda} - \lambda) = 0$, $\underline{\beta} = \beta$, $\underline{\sigma}^2 = \sigma^2$. That is, the limiting log-likelihood function is flat along the set of parameter values satisfying $\underline{\alpha} = \alpha - c_*\lambda + c_*\underline{\lambda}$. This implies that the typical identifiable uniqueness assumption maintained to prove consistency of an estimator does not hold at the true parameter values; compare, e.g., Gallant and White (1988, Chapter 3) and Pötscher and Prucha (1997, Chapter 3 and 4.6). This in turn suggests that also the maximum likelihood estimator, as well as any other estimator, will be inconsistent.

3. 2SLS and OLS in the case of panel data

3.1. The model

Our panel data generalization of the model in (1) is $(1 \le t \le T,\ N \ge 1)$

$$
\begin{aligned}
y_{tN} &= \alpha e_N + X_{tN}\beta + \lambda W_N y_{tN} + \varepsilon_{tN} \\
&= Z_{tN}\gamma + \varepsilon_{tN}
\end{aligned}
\tag{10}
$$

where $Z_{tN} = (e_N, X_{tN}, W_N y_{tN})$ and $\gamma = (\alpha, \beta', \lambda)'$. Here $y_{tN} = (y_{t1,N}, \dots, y_{tN,N})'$ is the $N \times 1$ vector of observations on the dependent variable in the t-th 'setting', X_{tN} is the $N \times k$ matrix of observations on the exogenous variables in the t-th 'setting', $\varepsilon_{tN} = (\varepsilon_{t1,N}, \dots, \varepsilon_{tN,N})'$ is the $N \times 1$ vector of disturbance terms in the t-th 'setting', and W_N is defined above in Assumption 3. These 'settings' could relate to time periods, or spatial entities for which there are multiple observations. Examples of such spatial entities would be villages, as in Case (1991), schools, etc. We again conditionalize our results on the realized values of the exogenous variables, and so we will take the matrices X_{tN} to be matrices of constants. We have assumed a balanced panel for ease of presentation; it will be clear that our results do not require a balanced panel. In the following we assume the number of settings T is greater than one. In our asymptotic analysis N is assumed to tend to infinity.

The model in (10) can be expressed as

$$\mathbf{y}_N = \mathbf{Z}_N\gamma + \boldsymbol{\varepsilon}_N \tag{11}$$

where

$$\mathbf{y}_N = [y'_{1N}, \ldots, y'_{TN}]',$$
$$\mathbf{Z}_N = [Z'_{1N}, \ldots, Z'_{TN}]',$$
$$\boldsymbol{\varepsilon}_N = [\varepsilon'_{1N}, \ldots, \varepsilon'_{TN}]'.$$

Let $D_{tN} = (e_N, X_{tN})$ for $t = 1, \ldots, T$. Our assumptions for the panel data model are:

Assumption 5. $|\lambda| < 1$. Also, for all $1 \leq t \leq T$, $1 \leq i \leq N$, and $N \geq 1$ the error terms $\varepsilon_{ti,N}$ are identically distributed with mean zero and finite variance σ^2. In addition, for each $N \geq 1$ the error terms $\varepsilon_{11,N}, \ldots, \varepsilon_{1N,N}, \ldots, \varepsilon_{T1,N}, \ldots, \varepsilon_{TN,N}$ are independently distributed.

Assumption 6. For each $t = 1, \ldots, T$ the elements of the sequence of matrices X_{tN} are uniformly bounded in absolute value and $Q_{DD,t} = \lim_{N\to\infty} N^{-1}D'_{tN}D_{tN}$ is finite and nonsingular.

It proves again convenient to decompose $Q_{DD,t}$ as

$$Q_{DD,t} = \begin{bmatrix} 1 & \mu'_{x,t} \\ \mu_{x,t} & Q_{XX,t} \end{bmatrix}.$$

Clearly $\mu_{x,t}$ is a $k \times 1$ vector, and $Q_{XX,t}$ is a $k \times k$ matrix. Let $\bar{x}_{tN} = N^{-1}X'_{tN}e_N$ be the vector of sample means of the exogenous regressors for the t-th setting. Then for future reference note that given the above assumption $\bar{x}_{tN} \to \mu_{x,t}$ and $N^{-1}X'_{tN}X_{tN} \to Q_{XX,t}$.

Observe that the panel data model in (10) corresponds to the model in (1) specified for each of the 'settings' $t = 1, \ldots, T$. Applying (4)–(6) we see that within each setting the spatial lag $W_N y_{tN}$ will be asymptotically proportional to the unit vector. However, unless $\mu_{x,t}$ is the same for each setting, the magnitude of the proportionality factor will vary across settings, and thus the spatial lag in the stacked model (11), that is $(I_T \otimes W_N)\mathbf{y}_N$, will not be proportional to the (stacked) unit vector. This suggests that the availability of panel data should alleviate the basic estimation problem discussed in Section 2 in the case of a single cross section.[5] Applying (4)–(6) and (9) to each 'setting' it is not difficult to see that

[5] It should also be clear, however, that panel data will not alleviate the problem if the intercept in (10) is allowed to vary with the 'setting' t, i.e. if the model has fixed effects.

$$\mathbf{Q}_{ZZ} = \lim_{N\to\infty} N^{-1}E(\mathbf{Z}_N)'E(\mathbf{Z}_N) = \sum_{t=1}^{T}\begin{bmatrix} 1 & \mu'_{x,t} & c_{*,t} \\ \mu_{x,t} & Q_{XX,t} & c_{*,t}\mu_{x,t} \\ c_{*,t} & c_{*,t}\mu_{x,t} & c^2_{*,t} \end{bmatrix} \tag{12}$$

with $c_{*,t} = (\alpha + \mu'_{x,t}\beta)/(1-\lambda)$. Our remaining assumption is as follows:

Assumption 7. The matrix $\mathbf{Q}_{ZZ}$ is nonsingular.

Clearly, as observed in the previous section, for $T = 1$ this assumption would be violated. However, an inspection of the expression for $\mathbf{Q}_{ZZ}$ shows that in general this matrix will not be singular, except in very special cases if $T > 1$. Those cases are ruled out by Assumption 7. In particular, Assumption 7 rules out the case in which $\beta = 0$, as well as the case in which $\mu_{x,t} = \mu_x$ for all $t = 1, \ldots, T$.

3.2. 2SLS and OLS estimation

The form of the instruments we use for the 2SLS estimator is partially based on the results obtained in Section 2. As remarked, the panel data model in (10) corresponds to the model in (1) specified for each of the 'settings' $t = 1, \ldots, T$. Applying (4) to the t-th 'setting' we see that the only deterministic regressors in the reduced form for $W_N y_{tN}$ are e_N and X_{tN}. The coefficient corresponding to e_N in the model solution for $W_N y_{tN}$ will clearly involve $\bar{x}_{tN}$, and therefore, consistent with our remarks above, will typically not be the same for each 'setting'. This suggests that the list of instruments underlying the 2SLS procedure should contain T dummy variables instead of just one constant term. For the purposes of generality we also interact these dummy variables with the regressor matrices X_{tN}, $t = 1, \ldots, T$. More specifically, the instrument matrix we consider is

$$\mathbf{H}_N = \begin{bmatrix} D_{1N} & 0 & . & . & 0 \\ 0 & . & & & . \\ . & & . & & . \\ . & & & . & 0 \\ 0 & . & . & 0 & D_{TN} \end{bmatrix} \tag{13}$$

Then, the 'second' stage regressor matrix is $\hat{\mathbf{Z}}_N = \mathbf{P}_{\mathbf{H}_N}\mathbf{Z}_N$, where $\mathbf{P}_{\mathbf{H}_N} = \mathbf{H}_N(\mathbf{H}'_N\mathbf{H}_N)^{-1}\mathbf{H}'_N$. It is not difficult to see that $\hat{\mathbf{Z}}_N$ can be expressed as

$$\hat{\mathbf{Z}}'_N = [\hat{Z}'_{1N}, \ldots, \hat{Z}'_{TN}] \tag{14}$$

with $\hat{Z}_{tN} = P_{D_{tN}}Z_{tN}$ and $P_{D_{tN}} = D_{tN}(D'_{tN}D_{tN})^{-1}D'_{tN}$. Since D_{tN} includes both e_N and X_{tN}, it follows that $P_{D_{tN}}e_N = e_N$ and $P_{D_{tN}}X_{tN} = X_{tN}$. Therefore

$$\hat{Z}_{tN} = (e_N, X_{tN}, P_{D_{tN}}W_N y_{tN}). \tag{15}$$

The 2SLS estimator of γ based on the panel data model in (14) is now given by

$$\tilde{\gamma}_{2\mathrm{SLS},N} = (\hat{\mathbf{Z}}'_N \hat{\mathbf{Z}}_N)^{-1} \hat{\mathbf{Z}}'_N \mathbf{y}_N. \tag{16}$$

The corresponding least squares estimator of γ is

$$\tilde{\gamma}_{\mathrm{OLS},N} = (\mathbf{Z}'_N \mathbf{Z}_N)^{-1} \mathbf{Z}'_N \mathbf{y}_N. \tag{17}$$

The next theorem, whose proof is given in Appendix A, establishes the asymptotic properties of the 2SLS and OLS estimator.

Theorem 1. *Given the model* (10) *and Assumptions* 3, 5–7

(a) $N^{1/2}(\tilde{\gamma}_{2\mathrm{SLS},N} - \gamma) \xrightarrow{D} N(0, \sigma^2 \mathbf{Q}_{ZZ}^{-1})$ as $N \to \infty$,
(b) $p\ \lim_{N\to\infty} N^{-1}(\hat{\mathbf{Z}}'_N \hat{\mathbf{Z}}_N) = \lim_{N\to\infty} N^{-1} E(\mathbf{Z}'_N) E(\mathbf{Z}_N) = \mathbf{Q}_{ZZ}$,
(c) $p\ \lim_{N\to\infty} N^{1/2}(\tilde{\gamma}_{2\mathrm{SLS},N} - \tilde{\gamma}_{\mathrm{OLS},N}) = 0$,
(d) $p\ \lim_{N\to\infty} \tilde{\sigma}_N^2 = \sigma^2$ where

$$\tilde{\sigma}_N^2 = (NT)^{-1}(\mathbf{y}_N - \mathbf{Z}_N \tilde{\gamma}_{2\mathrm{SLS},N})'(\mathbf{y}_N - \mathbf{Z}_N \tilde{\gamma}_{2\mathrm{SLS},N}).$$

A number of points should be noted about Theorem 1. Firstly, part (a) implies that the 2SLS estimator is consistent and asymptotically normal. Part (b) implies that the 2SLS estimator is the efficient instrumental variable estimator, since the limiting variance–covariance matrix is the inverse of the limit of the second moment of the mean of the regressor matrix. Part (c) implies that the 2SLS and OLS estimators are asymptotically equivalent. We note that in proving this result we establish, among other things, that

$$p \lim_{N\to\infty} N^{-1}(\hat{\mathbf{Z}}'_N \hat{\mathbf{Z}}_N) = p \lim_{N\to\infty} N^{-1} \mathbf{Z}_N \mathbf{Z}_N. \tag{18}$$

Part (d) implies that the typical estimator of the variance of the disturbance term is consistent.

Theorem 1 suggests that finite sample inferences concerning γ can be based on the approximation

$$\tilde{\gamma}_{2\mathrm{SLS},N} \cong N(\gamma, \tilde{\sigma}_N^2 (\hat{\mathbf{Z}}'_N \hat{\mathbf{Z}}_N)^{-1}),$$

or via (18)

$$\tilde{\gamma}_{2\mathrm{SLS},N} \cong N(\gamma, \tilde{\sigma}_N^2 (\mathbf{Z}'_N \mathbf{Z}_N)^{-1}).$$

In further research it may be of interest to analyze the small sample properties of the 2SLS and OLS estimator, and the quality of the approximate distribution derived from large sample theory, via a Monte Carlo study.

Acknowledgements

We thank the referees for helpful suggestions. We also gratefully acknowledge financial support from the National Science Foundation through grant SES-0001780.

Appendix A

Proof of Lemma 1

Let $\hat{\gamma}_{2SLS,N} = (\hat{\alpha}_{2SLS,N}, \hat{\beta}'_{2SLS,N}, \hat{\lambda}_{2SLS,N})'$. To prove Lemma 1 it suffices to show that $p\lim_{N\to\infty} \hat{\lambda}_{2SLS,N} \neq \lambda$. Let $P_{D_N} = D_N(D'_N D_N)^{-1}D'_N$. Then, by partitioned matrix inversion, $\hat{\lambda}_{2SLS,N}$ can be expressed as

$$\begin{aligned}\hat{\lambda}_{2SLS,N} &= \left([(I_N - P_{D_N})P_{H_N}W_N y_N]'[(I_N - P_{D_N})P_{H_N}W_N y_N]\right)^{-1} \\ &\quad \times [(I_N - P_{D_N})P_{H_N}W_N y_N]' y_N.\end{aligned} \tag{A.1}$$

Note that $(I_N - P_{D_N})D_N = 0$; also, since H_N contains the columns of D_N, $P_{H_N}D_N = D_N$ and so

$$P_{H_N}P_{D_N} = P_{D_N}. \tag{A.2}$$

Therefore, replacing y_N by its model expression in (1), and then using (A.2), it follows from (A.1) that

$$\begin{aligned}\hat{\lambda}_{2SLS,N} &= \lambda + [(W_N y_N)'(P_{H_N} - P_{D_N})(W_N y_N)]^{-1} \\ &\quad \times (W_N y_N)'(P_{H_N} - P_{D_N})\varepsilon_N \\ &= \lambda + \frac{(W_N y_N)'(P_{H_N} - P_{D_N})\varepsilon_N}{(W_N y_N)'(P_{H_N} - P_{D_N})(W_N y_N)}.\end{aligned} \tag{A.3}$$

In light of (4), $W_N y_N$ is linear in e_N, X_N, and ε_N. Since both H_N and D_N contain e_N and X_N it follows that $P_{H_N}e_N = e_N$, $P_{H_N}X_N = X_N$, $P_{D_N}e_N = e_N$, and $P_{D_N}X_N = X_N$. Therefore, from (4)

$$\begin{aligned}P_{H_N}W_N y_N &= c_N e_N + \phi_{2N}X_N\beta + \phi_{2N}P_{H_N}\varepsilon_N \\ P_{D_N}W_N y_N &= c_N e_N + \phi_{2N}X_N\beta + \phi_{2N}P_{D_N}\varepsilon_N\end{aligned} \tag{A.4}$$

and so the numerator in (A.3) is

$$(W_N y_N)'(P_{H_N} - P_{D_N})\varepsilon_N = \phi_{2N}[\varepsilon'_N P_{H_N}\varepsilon_N - \varepsilon'_N P_{D_N}\varepsilon_N]. \tag{A.5}$$

The result in (4) and (A.4) imply that the denominator in (A.3) is expressible as

$$\begin{aligned}(W_N y_N)'[P_{H_N} - P_{D_N}](W_N y_N) &= \phi_{2N}(W_N y_N)'[P_{H_N} - P_{D_N}]\varepsilon_N \\ &= \phi^2_{2N}[\varepsilon'_N P_{H_N}\varepsilon_N - \varepsilon'_N P_{D_N}\varepsilon_N].\end{aligned} \tag{A.6}$$

It follows from (A.3), (A.5) and (A.6) that

$$\hat{\lambda}_{2SLS,N} = \lambda + \frac{1}{\phi_{2N}} \tag{A.7}$$

and so via (5)

$$p \lim_{N\to\infty} \left|\hat{\lambda}_{2SLS,N}\right| = \infty \neq |\lambda|. \tag{A.8}$$

Proof of Lemma 2

Let $\hat{\gamma}'_{OLS,N} = (\hat{\alpha}_{OLS,N}, \hat{\beta}'_{OLS,N}, \hat{\lambda}_{OLS,N})$. To prove Lemma 2 it suffices to show that $p \lim_{N\to\infty} \hat{\lambda}_{OLS,N} \neq \lambda$. By partitioned matrix inversion the model in (1) implies

$$\begin{aligned}\hat{\lambda}_{OLS,N} &= \left(\left[(I_N - P_{D_N})W_N y_N\right]'\left[(I_N - P_{D_N})W_N y_N\right]\right)^{-1} \\ &\quad \times \left[(I_N - P_{D_N})W_N y_N\right]' y_N \\ &= \lambda + \frac{(W_N y_N)'(I_N - P_{D_N})\varepsilon_N}{(W_N y_N)'(I_N - P_{D_N})W_N y_N}.\end{aligned} \tag{A.9}$$

Recalling that $P_{D_N} e_N = e_N$ and $P_{D_N} X_N = X_N$, the results in (4) and (A.4) imply

$$\hat{\lambda}_{OLS,N} = \lambda + \frac{\phi_{2N}(\varepsilon'_N\varepsilon_N - \varepsilon'_N P_{D_N}\varepsilon_N)}{\phi^2_{2N}(\varepsilon'_N\varepsilon_N - \varepsilon'_N P_{D_N}\varepsilon_N)} = \lambda + \frac{1}{\phi_{2N}} \tag{A.10}$$

and so via (5)

$$p \lim_{N\to\infty} \left|\hat{\lambda}_{OLS,N}\right| = \infty \neq |\lambda|.$$

As a point of interest we note that in light of (A.7) and (A.10) we have $\hat{\lambda}_{2SLS,N} = \hat{\lambda}_{OLS,N}$. The equality between the 2SLS and OLS estimators for λ does not extend, however, to all components of the parameter vector γ.

Proof of Theorem 1

Parts (a) and (b): Since $\hat{\mathbf{Z}}'_N\hat{\mathbf{Z}}_N = \hat{\mathbf{Z}}'_N\mathbf{Z}_N$, the model in (11) implies

$$\begin{aligned}N^{1/2}(\tilde{\gamma}_{2SLS,N} - \gamma) &= \left(N^{-1}\hat{\mathbf{Z}}'_N\hat{\mathbf{Z}}_N\right)^{-1}N^{-1/2}\hat{\mathbf{Z}}'_N\boldsymbol{\varepsilon}_N \\ &= \left(N^{-1}\sum_{t=1}^{T}\hat{Z}'_{tN}\hat{Z}_{tN}\right)^{-1}N^{-1/2}\sum_{t=1}^{T}\hat{Z}'_{tN}\varepsilon_{tN}.\end{aligned} \tag{A.11}$$

Consider the first term on the r.h.s. of (A.11). Extending the result in (4) to the panel data model in (11), and noting that $P_{D_{tN}}D_{tN} = D_{tN}$ it follows from (15) that

$$\hat{Z}_{tN} = (D_{tN}, c_{tN}e_N + \Pi_{tN}) \tag{A.12}$$

where

$$
\begin{aligned}
c_{tN} &= N\phi_{1N}[\alpha + \bar{x}'_{tN}\beta + \bar{\varepsilon}_{tN}] + \alpha\phi_{2N}, \\
\Pi_{tN} &= \phi_{2N}X_{tN}\beta + \phi_{2N}P_{D_{tN}}\varepsilon_{tN}, \\
\bar{\varepsilon}_{tN} &= N^{-1}e'_N\varepsilon_{tN}.
\end{aligned}
$$

The results in (A.12) imply

$$
N^{-1}\hat{Z}'_{tN}\hat{Z}_{tN} = \begin{bmatrix} N^{-1}D'_{tN}D_{tN} & c_{tN}N^{-1}D'_{tN}e_N + N^{-1}D'_{tN}\Pi_{tN} \\ c_{tN}N^{-1}e'_ND_{tN} + N^{-1}\Pi'_{tN}D_{tN} & c^2_{tN} + N^{-1}\Pi'_{tN}\Pi_{tN} + 2c_{tN}N^{-1}e'_{tN}\Pi_{tN} \end{bmatrix}. \tag{A.13}
$$

Consider the elements of the matrix in (A.13). As a preliminary observation we note that applying the results (4)–(6) to the t-th setting gives

$$
c_{tN} \to c_{*,t} = \frac{\alpha + \mu'_{x,t}\beta}{1-\lambda} < \infty
$$

as $N \to \infty$. Now let $\Psi_{tN} = N^{-1}D'_{tN}\varepsilon_{tN}$. Then, Assumptions 5 and 6 imply $E(\Psi_{tN}) = 0$ and $E(\Psi_{tN}\Psi'_{tN}) = \sigma^2N^{-2}D'_{tN}D_{tN} \to 0$ as $N \to \infty$, and so, via Chebyshev's inequality, $p\ \lim_{N\to\infty} N^{-1}D'_{tN}\varepsilon_{tN} = 0$. As a consequence

$$
\begin{aligned}
&N^{-1}D'_{tN}\Pi_{tN} = \phi_{2N}(N^{-1}D'_{tN}X_{tN})\beta + \phi_{2N}(N^{-1}D'_{tN}\varepsilon_{tN}) \xrightarrow{P} 0 \\
&N^{-1}\Pi'_{tN}\Pi_{tN} = \phi^2_{2N}\beta'(N^{-1}X'_{tN}X_{tN})\beta \\
&\quad + \phi^2_{2N}(N^{-1}\varepsilon'_{tN}D_{tN})(N^{-1}D'_{tN}D_{tN})^{-1}(N^{-1}D'_{tN}\varepsilon_{tN}) \\
&\quad + 2\phi^2_{2N}\beta'(N^{-1}X'_{tN}D_{tN})(N^{-1}D'_{tN}D_{tN})^{-1}(N^{-1}D'_{tN}\varepsilon_{tN}) \\
&\quad \xrightarrow{P} 0
\end{aligned} \tag{A.14}
$$

as $N \to \infty$. Since D_{tN} contains e_N we have furthermore

$$
c_{tN}N^{-1}e'_N\Pi_{tN} \xrightarrow{P} 0 \tag{A.15}
$$

as $N \to \infty$.

The results in (A.13)–(A.15), and Assumption 6 imply that for $t = 1, \ldots, T$

$$
\begin{aligned}
p\lim_{N\to\infty} N^{-1}\hat{Z}'_{tN}Z_{tN} &= p\lim_{N\to\infty}\begin{bmatrix} N^{-1}D'_{tN}D_{tN} & c_{tN}N^{-1}D'_{tN}e_N \\ c_{tN}N^{-1}e'_ND_{tN} & c^2_{tN} \end{bmatrix} \\
&= \begin{bmatrix} 1 & \mu'_{x,t} & c_{*,t} \\ \mu_{x,t} & Q_{XX,t} & c_{*,t}\mu_{x,t} \\ c_{*,t} & c_{*,t}\mu'_{*,t} & c^2_{*,t} \end{bmatrix}
\end{aligned} \tag{A.16}
$$

and consequently

$$p\lim_{N\to\infty} N^{-1}\hat{\mathbf{Z}}'_N\hat{\mathbf{Z}}_N = \sum_{t=1}^{T}\begin{bmatrix} 1 & \mu'_{x,t} & c_{*,t} \\ \mu_{x,t} & Q_{XX,t} & c_{*,t}\mu_{x,t} \\ c_{*,t} & c_{*,t}\mu'_{x,t} & c^2_{*,t} \end{bmatrix} = \mathbf{Q}_{ZZ}. \tag{A.17}$$

In light of Assumptions 6 and 7 this matrix is finite and nonsingular.

Now consider the second term on the r.h.s. of (A.11). First note from (A.12) that

$$\begin{aligned} N^{-1/2}\Pi'_{tN}\varepsilon_{tN} &= \left(N^{1/2}\phi_{2N}\right)\beta'\left(N^{-1}X'_{tN}\varepsilon_{tN}\right) \\ &\quad + \left(N^{1/2}\phi_{2N}\right)\left(N^{-1}\varepsilon'_{tN}D_{tN}\right)\left(N^{-1}D'_{tN}D_{tN}\right)^{-1}\left(N^{-1}D'_{tN}\varepsilon_{tN}\right). \end{aligned} \tag{A.18}$$

From (3) and (4) observe that $N^{1/2}\phi_{2N}\to 0$. Since $p\lim_{N\to\infty} N^{-1}D'_{tN}\varepsilon_{tN}=0$, and utilizing Assumption 6 it follows that

$$p\lim_{N\to\infty} N^{-1/2}\Pi'_{tN}\varepsilon_{tN} = 0, \quad t=1,\ldots,T. \tag{A.19}$$

Also note that

$$p\lim_{N\to\infty} (c_{N,t}-c_{*,t})N^{-1/2}e'_N\varepsilon_{tN} = 0, \quad t=1,\ldots,T \tag{A.20}$$

since $p\lim_{N\to\infty}(c_{N,t}-c_{*,t})=0$ and $N^{-1/2}e'_N\varepsilon_{tN}\xrightarrow{D}N(0,\sigma^2_\epsilon)$ by the Lindeberg–Lévy central limit theorem.

Now define

$$\overline{\mathbf{Z}}_N = \left(\overline{Z}'_{1N},\ldots,\overline{Z}'_{TN}\right)'$$

where

$$\overline{Z}_{tN} = (D_{tN}, c_{*,t}e_N), \quad t=1,\ldots,T.$$

Then clearly

$$\lim_{N\to\infty} N^{-1}\overline{\mathbf{Z}}'_N\overline{\mathbf{Z}}_N = \mathbf{Q}_{ZZ}, \tag{A.21}$$

which is, as remarked above, finite and nonsingular. Assumptions 6 and 7 imply that the elements of $\overline{\mathbf{Z}}_N$ are uniformly bounded in absolute value (observing that T is fixed). In light of Assumption 5 all conditions for the central limit theorem given in Kelejian and Prucha (1998, p. 112) are thus seen to be satisfied. Applying this theorem yields

$$N^{-1/2}\overline{\mathbf{Z}}'_N\boldsymbol{\varepsilon}_N \xrightarrow{D} N\left(0,\sigma^2\mathbf{Q}_{ZZ}\right). \tag{A.22}$$

From (A.12), (A.19) and (A.20) it is readily seen that

$$p\lim_{N\to\infty}\left(N^{-1/2}\hat{\mathbf{Z}}'_N\boldsymbol{\varepsilon}_N - N^{-1/2}\overline{\mathbf{Z}}'_N\boldsymbol{\varepsilon}_N\right) = 0. \tag{A.23}$$

Consequently we also have

$$N^{-1/2}\hat{\mathbf{Z}}'_N\boldsymbol{\varepsilon}_N \xrightarrow{D} N(0, \sigma^2\mathbf{Q}_{ZZ}) \tag{A.24}$$

as $N\to\infty$. Parts (a) and (b) of Theorem 1 now follows from (A.11), (A.17), (A.24) and (12).

Part (c): From (11) and (17) we have

$$N^{1/2}(\tilde{\gamma}_{\mathrm{OLS},N} - \gamma) = (N^{-1}\mathbf{Z}'_N\mathbf{Z}_N)^{-1}N^{-1/2}\mathbf{Z}'_N\boldsymbol{\varepsilon}_N. \tag{A.25}$$

Consider $N^{-1}(\mathbf{Z}'_N\mathbf{Z}_N)$. Note again from the extension of (4)–(6) that

$$Z_{tN} = (D_{tN}, c_{tN}e_N + \phi_{2N}X_{tN}\beta + \phi_{2N}\varepsilon_{tN}).$$

Since Assumptions 5 and 6 imply that $p\lim_{N\to\infty} \phi_{2N}N^{-1}D'_{tN}D_{tN} = 0$ and $p\lim_{N\to\infty} N^{-1}D'_{tN}\varepsilon_{tN} = 0$ it should be clear that

$$p\lim_{N\to\infty} N^{-1}\hat{\mathbf{Z}}'_N\hat{\mathbf{Z}}_N = p\lim_{N\to\infty} N^{-1}\mathbf{Z}'_N\mathbf{Z}_N = \mathbf{Q}_{ZZ}. \tag{A.26}$$

Therefore the first factors in (A.11) and (A.25) have the same probability limit.

Now consider the second factors in (A.11) and (A.25). Given (A.19), recalling that $N^{1/2}\phi_{2N}\to 0$ as $N\to\infty$, and then noting that Assumptions 5 and 6 imply

$$\begin{aligned} &p\lim_{N\to\infty} (N^{1/2}\phi_{2N})N^{-1}\varepsilon'_{tN}\varepsilon_{tN} = 0,\\ &p\lim_{N\to\infty} (N^{1/2}\phi_{2N})N^{-1}X'_{tN}\varepsilon_{tN} = 0, \end{aligned} \tag{A.27}$$

it follows that

$$p\lim_{N\to\infty} (N^{-1/2}\mathbf{Z}'_N\boldsymbol{\varepsilon}_N - N^{-1/2}\hat{\mathbf{Z}}'_N\boldsymbol{\varepsilon}_N) = 0. \tag{A.28}$$

Part (c) follows from (A.26) and (A.28).

Part (d): Let

$$\begin{aligned} \tilde{\boldsymbol{\varepsilon}}_N &= \mathbf{y}_N - \mathbf{Z}_N\tilde{\gamma}_{2\mathrm{SLS},N}\\ &= \boldsymbol{\varepsilon}_N - \mathbf{Z}_N\Delta_N \end{aligned}$$

where $\Delta_N = \tilde{\gamma}_{2\mathrm{SLS},N} - \gamma$. Then

$$\tilde{\sigma}^2_N = (NT)^{-1}\boldsymbol{\varepsilon}'_N\boldsymbol{\varepsilon}_N - 2(NT)^{-1}\boldsymbol{\varepsilon}'_N\mathbf{Z}_N\Delta_N + \Delta'_N(NT)^{-1}\mathbf{Z}'_N\mathbf{Z}_N\Delta_N. \tag{A.29}$$

For a given value of T, Assumption 5 and Khintchine's law of large numbers, imply

$$p\lim_{N\to\infty} (NT)^{-1}\boldsymbol{\varepsilon}'_N\boldsymbol{\varepsilon}_N = \sigma^2. \tag{A.30}$$

The consistency of $\tilde{\gamma}_{2\mathrm{SLS},N}$ implies $p\lim_{N\to\infty} \Delta_N = 0$. Given this, (A.24), (A.26) and (A.28)

$$p\lim_{N\to\infty} (NT)^{-1}\boldsymbol{\varepsilon}_N'\mathbf{Z}_N\Delta_N = 0,$$
$$p\lim_{N\to\infty} (NT)^{-1}\Delta_N'\mathbf{Z}_N'\mathbf{Z}_N\Delta_N = 0. \tag{A.31}$$

Part (d) then follows from (A.30) and (A.31).

References

Amemiya, T., 1985. Advanced Econometrics. Harvard University Press, Cambridge.

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

Anselin, L., Bera, A., Florax, R., Yoon, M., 1996. Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics 26, 77–104.

Boarnet, M., 1994. An empirical model of intrametropolitan population and employment growth. Papers in Regional Science 73, 135–152.

Bollinger, C., Ihlanfeldt, K., 1997. The impact of rapid rail transit on economic development: the case of Atlanta's Marta. Journal of Urban Economics 42, 179–204.

Case, A., 1991. Spatial patterns in household demand. Econometrica 59, 953–966.

Case, A., 1992. Neighborhood influence and technological change. Regional Science 22, 491–508.

Cliff, A., Ord, J., 1973. Spatial Autocorrelation. Pion, London.

Cliff, A., Ord, J., 1981. Spatial Processes, Models and Applications. Pion, London.

Conley, T., 1999. GMM estimation with cross sectional dependence. Journal of Econometrics 92, 1–45.

Cressie, N.A.C., 1993. Statistics of Spatial Data. Wiley, New York.

Gallant, A.R., 1987. Nonlinear Statistical Models. Wiley, New York.

Gallant, A.R., White, H., 1988. A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models. Basil Blackwell, New York.

Harville, D.A., 1988. Mixed-model methodology: theoretical justifications and future directions. In: American Statistical Association, Proceedings of the Statistical Computing Section, pp. 41–49.

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

Kelejian, H.H., Prucha, I.R., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review 40, 509–533.

Kelejian, H.H., Prucha, I.R., 2001. On the asymptotic distribution of the Moran I test statistic with applications. Journal of Econometrics 104, 219–257.

Lee, L.F., 1999a. Best Spatial Two-stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances. Department of Economics, HKUST, Hong Kong.

Lee, L.F., 1999b. Consistency and Efficiency of Least Squares Estimation for Mixed Regressive, Spatial Autoregessive Models. Department of Economics, HKUST, Hong Kong.

Lee, L.F., 1999c. Asymptotic Distributions of Maximum Likelihood Estimators for Spatial Autoregressive Models. Department of Economics, HKUST, Hong Kong.

Pace, R., Barry, R., 1977. Sparse spatial autoregressions. Statistics and Probability Letters 33, 291–297.

Pötscher, B.M., Prucha, I.R., 1997. Dynamic Nonlinear Econometric Models, Asymptotic Theory. Springer Verlag, New York.

Rey, S.J., Boarnet, M.G., 1998. A Taxonomy of Spatial Econometric Models for Simultaneous Equation Systems. Department of Geography, San Diego State University.

Splitstoser, G., 1999. A Spatial Model for Legislator Ideology. Department of Economics, University of Maryland, Term paper.