Econometric Theory, **26**, 2010, 301–310.
doi:10.1017/S0266466609090756

NOTES AND PROBLEMS

SPECIFICATION OF VARIANCE MATRICES FOR PANEL DATA MODELS

JAN R. MAGNUS AND CHRIS MURIS
Tilburg University

Many regression models have two dimensions, say time ($t = 1, \ldots, T$) and households ($i = 1, \ldots, N$), as in panel data, error components, or spatial econometrics. In estimating such models we need to specify the structure of the error variance matrix Ω, which is of dimension $TN \times TN$. If TN is large, then direct computation of the determinant and inverse of Ω is not practical. In this note we define structures of Ω that allow the computation of its determinant and inverse, only using matrices of orders T and N, and at the same time allowing for heteroskedasticity, for household- or station-specific autocorrelation, and for time-specific spatial correlation.

1. INTRODUCTION

We consider regression models with two dimensions, which we denote by T (say time) and N (say households or stations), such as

$$y_{it} = f_i(X_{it}, \beta_i) + u_{it} \qquad (i = 1, \ldots, N;\ t = 1, \ldots, T).$$

In estimating such models we need to specify the structure of the variance matrix Ω of the errors u_{it}, which will be of dimension $TN \times TN$. We shall assume that TN is so large that direct computation of its determinant and inverse is not practical. Thus we need to find a structure of Ω that allows the computation of its determinant and inverse, only using matrices of orders T and N but not TN. In this note we attempt to obtain maximum flexibility of the variance matrix under precisely this constraint. The flexibility that we aim for should allow for heteroskedasticity in the errors, for household- or station-specific autocorrelation, and also for time-specific spatial correlation.

Problems of this nature arise in the panel data and error components literature; see Baltagi and Raj (1992), Baltagi (2001), and Arellano (2003) for useful reviews and historical details. They are also important in the closely related area of spatial econometrics; see Anselin (1988), Anselin and Bera (1998), Driscoll and Kraay

We are grateful to Paolo Paruolo and two referees for their helpful and constructive comments. Address correspondence to Jan R. Magnus, Department of Econometrics and Operations Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands; e-mail: magnus@uvt.nl.

 0266-4666/10 $15.00

(1998), Baltagi, Song, and Koh (2003), Baltagi, Song, Jung, and Koh (2007), and Kapoor, Kelejian, and Prucha (2007). The idea of introducing heteroskedasticity into error component models is discussed in Baltagi and Griffin (1988) and Li and Stengos (1994). Closest to our approach are the papers by Searle and Henderson (1979) and Wansbeek and Kapteyn (1982); the authors try to understand, like us, which class of variance matrices is appropriate for models with two dimensions.

We combine the errors in a matrix

$$U = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1T} \\ u_{21} & u_{22} & \dots & u_{2T} \\ \vdots & \vdots & & \vdots \\ u_{N1} & u_{N2} & \dots & u_{NT} \end{pmatrix}, \tag{1}$$

and we define its T columns and N rows as

$$U = (u_1, u_2, \dots, u_T), \qquad U' = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_N). \tag{2}$$

Letting $u = \text{vec}(U)$, we then have

$$\Omega = \text{var}(u) = \begin{pmatrix} \Omega_{11} & \Omega_{12} & \dots & \Omega_{1T} \\ \Omega_{21} & \Omega_{22} & \dots & \Omega_{2T} \\ \vdots & \vdots & & \vdots \\ \Omega_{T1} & \Omega_{T2} & \dots & \Omega_{TT} \end{pmatrix}, \tag{3}$$

where each of the submatrices is of order $N \times N$.

The simplest case is of course $\Omega = A \otimes B$, but this is usually not sufficiently general. The following result is often useful and is stated here separately because of its importance and also because we will refer to it in what follows. It is a special case of Lemma 2.1 of Magnus (1982).

LEMMA 1. *Let A_k ($k = 1, 2, \dots, K$) be symmetric idempotent matrices of order $T \times T$ and rank r_k satisfying $\sum_k A_k = I_T$ and let B_k ($k = 1, 2, \dots, K$) be positive definite of order $N \times N$. Define $\Omega = \sum_k (A_k \otimes B_k)$ of order $TN \times TN$. Then, Ω is positive definite, and its eigenvalues are the eigenvalues of $B_1, B_2, \dots, B_K$ with multiplicities $r_1, r_2, \dots, r_K$. Further,*

$$|\Omega| = \prod_{k=1}^{K} |B_k|^{r_k}, \qquad \Omega^{-1} = \sum_{k=1}^{K} (A_k \otimes B_k^{-1}).$$

In this note we emphasize structures like $\Omega = A_1 \otimes B_1 + A_2 \otimes B_2$, because a Kronecker-type structure seems natural in the context of two dimensions (T and N). Of course other structures can be defined that cannot be written in this way but still allow the computation of the determinant and inverse of Ω using only matrices of orders T and N but not TN.

This note is organized as follows. In Section 2 we make the simplifying assumption that the T variance matrices Ω_{tt} are free but that the correlation matrices are constant. In Section 3 we consider the two error components model, which is perhaps the main tool for panel data. We will see that considerably more flexibility is possible than previously utilized in the panel data literature. In Section 4 we study the case where assumptions on the columns of U are combined with an independence assumption on (linear combinations of) the rows. In Section 5 we consider the three error components model and try to understand why this more general setup does not lead to a more general specification. Section 6 concludes.

2. CONSTANT CORRELATION

Let us write the variance matrix Ω in terms of its correlation matrices

$$P_{st} = \Omega_{ss}^{-1/2}\Omega_{st}\Omega_{tt}^{-1/2},$$

so that Ω takes the form

$$\begin{pmatrix} \Omega_{11}^{1/2} & 0 & \dots & 0 \\ 0 & \Omega_{22}^{1/2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \Omega_{TT}^{1/2} \end{pmatrix} \begin{pmatrix} I_N & P_{12} & \dots & P_{1T} \\ P_{21} & I_N & \dots & P_{21} \\ \vdots & \vdots & & \vdots \\ P_{T1} & P_{T2} & \dots & I_N \end{pmatrix} \begin{pmatrix} \Omega_{11}^{1/2} & 0 & \dots & 0 \\ 0 & \Omega_{22}^{1/2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \Omega_{TT}^{1/2} \end{pmatrix}.$$

Suppose we like to keep maximum flexibility on the structure of the variance matrices Ω_{tt} but that we are willing to assume that all correlation matrices are the same: $P_{st} = P$. In the special case where $P = 0$, this means that the error vectors $\{u_t\}$ are uncorrelated. The current assumption is more general. Also, with a obvious change of indexes, we may assume that there is zero or constant correlation over households rather than over time.

The determinant and inverse of Ω can then be obtained from the following theorem.

THEOREM 1. *Let the st-th block of Ω be defined as*

$$\Omega_{st} = \begin{cases} B_t & \text{if } s = t, \\ B_s^{1/2} P B_t^{1/2} & \text{if } s \neq t, \end{cases}$$

where the B_t are all positive definite $N \times N$ matrices and P is a symmetric $N \times N$ matrix whose eigenvalues are bounded by

$$\lambda_{min}(P) > \frac{-1}{T-1}, \qquad \lambda_{max}(P) < 1.$$

Define the two matrices

$$C_1 = I_N + (T-1)P, \qquad C_2 = I_N - P.$$

Then Ω is positive definite, its determinant is given by

$$|\Omega| = |C_1||C_2|^{T-1} \prod_{t=1}^{T} |B_t|,$$

and the st-th block of Ω^{-1} is given by

$$\Omega^{st} = \begin{cases} \frac{1}{T} B_t^{-1/2}(C_1^{-1} + (T-1)C_2^{-1})B_t^{-1/2} & \text{if } s = t, \\ \frac{1}{T} B_s^{-1/2}(C_1^{-1} - C_2^{-1})B_t^{-1/2} & \text{if } s \neq t. \end{cases}$$

Proof. The matrix C_1 is positive definite because $\lambda_{min}(P) > -1/(T-1)$, and C_2 is positive definite because $\lambda_{max}(P) < 1$. Next, let

$$\bar{B} = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & B_T \end{pmatrix}, \qquad \Omega_1 = \begin{pmatrix} I_N & P & \dots & P \\ P & I_N & \dots & P \\ \vdots & \vdots & & \vdots \\ P & P & \dots & I_N \end{pmatrix},$$

so that $\Omega = \bar{B}^{1/2}\Omega_1\bar{B}^{1/2}$. Letting $p = \imath/\sqrt{T}$, where $\imath$ denotes the vector of ones, we can write Ω_1 more conveniently as

$$\Omega_1 = pp' \otimes C_1 + (I_T - pp') \otimes C_2.$$

Because pp' and $I_T - pp'$ are idempotent and sum to I_T, and because their ranks are 1 and $T-1$, respectively, we see from Lemma 1 that Ω_1 (and hence Ω) is positive definite and

$$|\Omega_1| = |C_1||C_2|^{T-1}, \qquad \Omega_1^{-1} = pp' \otimes C_1^{-1} + (I_T - pp') \otimes C_2^{-1}.$$

This implies that

$$|\Omega| = |\Omega_1||\bar{B}| = |C_1||C_2|^{T-1} \prod_{t=1}^{T} |B_t|$$

and

$$\begin{aligned} \Omega^{-1} &= \bar{B}^{-1/2}\Omega_1^{-1}\bar{B}^{-1/2} = \bar{B}^{-1/2}\big(pp' \otimes C_1^{-1} + (I_T - pp') \otimes C_2^{-1}\big)\bar{B}^{-1/2} \\ &= \bar{B}^{-1/2}\left(I_T \otimes C_2^{-1} + \frac{\imath\imath'}{T} \otimes \left(C_1^{-1} - C_2^{-1}\right)\right)\bar{B}^{-1/2}. \end{aligned}$$

The result follows. ■

3. TWO ERROR COMPONENTS

Although it is trivial to find the determinant and inverse of a simple Kronecker product $A \otimes B$, it is less trivial to do the same for the sum of two Kronecker products $A_1 \otimes B_1 + A_2 \otimes B_2$. In the special case where both A_2 and B_2 are positive definite we obtain Theorem 2.

THEOREM 2. *Let A_1 $(T \times T)$ and B_1 $(N \times N)$ be positive semidefinite matrices of rank $r_A \leq T$ and $r_B \leq N$, respectively. Write $A_1 = \Gamma\Gamma'$ and $B_1 = \Delta\Delta'$ where Γ and Δ have full column rank. Let A_2 and B_2 be positive definite matrices of order $T \times T$ and $N \times N$, respectively. Then the $TN \times TN$ matrix*

$$\Omega = \Gamma\Gamma' \otimes \Delta\Delta' + A_2 \otimes B_2$$

is positive definite with

$$|\Omega| = |A_2|^N |B_2|^T |\Omega_r|$$

and

$$\Omega^{-1} = A_2^{-1} \otimes B_2^{-1} - \left(A_2^{-1}\Gamma \otimes B_2^{-1}\Delta\right)\Omega_r^{-1}\left(\Gamma' A_2^{-1} \otimes \Delta' B_2^{-1}\right),$$

where

$$\Omega_r = I_{r_A} \otimes I_{r_B} + \Gamma' A_2^{-1}\Gamma \otimes \Delta' B_2^{-1}\Delta.$$

Proof. Write $\Omega = LL' + \Omega_2$ where Ω_2 is positive definite and L has full column rank, say r. Then,

$$|\Omega| = |\Omega_2|\,|I_r + L'\Omega_2^{-1}L|$$

and

$$\Omega^{-1} = \Omega_2^{-1} - \Omega_2^{-1}L\left(I_r + L'\Omega_2^{-1}L\right)^{-1}L'\Omega_2^{-1}.$$

Now let $\Omega_2 = A_2 \otimes B_2$ and $L = \Gamma \otimes \Delta$, and the results follow. ■

We notice that the determinant and inverse of Ω in Theorem 2 involve the determinant and the inverse of Ω_r. This matrix is a sum of two Kronecker products, like Ω, but the constituent matrices are all of full rank, unlike Ω, so that the expressions cannot be simplified further. This case therefore requires the calculation of the determinant and inverse of an $r_A r_B \times r_A r_B$ matrix instead of a $TN \times TN$ matrix, which may or may not be a sizable gain. Theorem 2 also shows that unless sufficient structure is imposed on A_1 and B_1 the size of the constituent matrices cannot be reduced to merely $T \times T$ and $N \times N$.

One case where the size of the constituent matrices *can* be reduced to merely $T \times T$ and $N \times N$, and which is therefore of particular interest, is the case where $r_A = 1$.

THEOREM 3. *Let A be a positive definite $T \times T$ matrix, a a nonzero $T \times 1$ vector, B_1 a positive semidefinite $N \times N$ matrix (possibly the null matrix), and B_2 a positive definite $N \times N$ matrix. Then the $TN \times TN$ matrix*

$$\Omega = aa' \otimes B_1 + A \otimes B_2$$

is positive definite with

$$|\Omega| = |A|^N |B_2|^{T-1} |B_2 + (a'A^{-1}a)B_1|$$

and

$$\Omega^{-1} = \frac{A^{-1}aa'A^{-1}}{a'A^{-1}a} \otimes \left(B_2 + (a'A^{-1}a)B_1\right)^{-1} + \left(A^{-1} - \frac{A^{-1}aa'A^{-1}}{a'A^{-1}a}\right) \otimes B_2^{-1}.$$

Proof. The theorem can be obtained as a special case of Theorem 2, but the following proof provides additional insight. Let

$$\alpha^2 = a'A^{-1}a, \qquad p = A^{-1/2}a/\alpha, \qquad C = B_2 + \alpha^2 B_1.$$

Then,

$$\Omega = aa' \otimes B_1 + A \otimes B_2 = (A^{1/2} \otimes I_N)\Omega_1(A^{1/2} \otimes I_N),$$

where

$$\begin{aligned}\Omega_1 &= A^{-1/2}aa'A^{-1/2} \otimes B_1 + I_T \otimes B_2 \\ &= pp' \otimes \alpha^2 B_1 + I_T \otimes B_2 \\ &= pp' \otimes C + (I_T - pp') \otimes B_2.\end{aligned}$$

Because pp' and $I_T - pp'$ are idempotent matrices that sum to I_T, and B_2 and C are both positive definite, it follows from Lemma 1 that Ω_1 is positive definite and that

$$|\Omega_1| = |B_2|^{T-1}|C|, \qquad \Omega_1^{-1} = pp' \otimes C^{-1} + (I_T - pp') \otimes B_2^{-1}.$$

The determinant and inverse of Ω now follow easily. ■

Special cases of Theorem 3 have been studied in the literature. Thus, Baltagi et al. (2003) consider the case where $a = \imath$, $B_1 = \sigma_1^2 I_N$, $A = \sigma_2^2 I_T$, and $B_2 = ((I_N - \lambda W)'(I_N - \lambda W))^{-1}$, which arises from a structure with random regional effects and spatially correlated errors. In particular, it is assumed that

$$u_{it} = v_i + \epsilon_{it}, \qquad \epsilon_t = \lambda W \epsilon_t + e_t,$$

where $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \ldots, \epsilon_{Nt})'$, v_i and ϵ_{it} are independent, the v_i are independent and identically distributed with common variance σ_1^2, and $\mathrm{var}(e_t) = \sigma_2^2 I_N$. In a sequel paper, Baltagi et al. (2007) assume in addition that the remainder term e_t follows a first-order autoregressive process by defining A to be the familiar AR(1) variance matrix. Theorems 2 and 3 show that considerably more generality is still feasible. By writing

$$u_{it} = \alpha_t v_i + \epsilon_{it},$$

and letting $a = (\alpha_1, \alpha_2, \dots, \alpha_T)'$, Theorem 3 allows complete freedom of choice (including singularity) over $B_1 = \text{var}(v_1, v_2, \dots, v_N)$ and only requires that $\text{cov}(\epsilon_s, \epsilon_t) = a_{st} B_2$ where $A = (a_{st})$ and B_2 are both positive definite. This allows the first error component to depend on t and generalizes the spatial structure in B_2 and the correlation structure in A. Application of Theorem 4, which follows, allows further generalization of the first error component.

4. WEAK ROW INDEPENDENCE

Clearly, some structure must be assumed to get a manageable variance matrix Ω. One often wants to make assumptions on the columns of the error matrix U, whereas it is reasonable to assume independence of (some transformation of) the rows. The following result is somewhat related to the work of Kapoor et al. (2007), who also wish to combine household- or station-specific autocorrelation with time-specific spatial correlation.

THEOREM 4. *Let the error vectors $u_1, u_2, \dots, u_T$ be generated by $u_t = B_t \epsilon_t$, where $\epsilon_1, \epsilon_2, \dots, \epsilon_T$ are the T columns of an $N \times T$ matrix E whose N rows are defined as $\tilde{\epsilon}_1', \tilde{\epsilon}_2', \dots, \tilde{\epsilon}_N'$, and assume that the vectors $\tilde{\epsilon}_1, \tilde{\epsilon}_2, \dots, \tilde{\epsilon}_N$ are independently distributed with mean zero and variance* $\text{var}(\tilde{\epsilon}_i) = A_i$. *Then,*

$$|\Omega| = \prod_{t=1}^{T} |B_t|^2 \prod_{i=1}^{N} |A_i|,$$

and, if all A_i and B_t are nonsingular, Ω is positive definite and the st-th block of Ω^{-1} is given by

$$\Omega^{st} = (B_s')^{-1} \text{diag}\left(A_1^{st}, A_2^{st}, \dots, A_N^{st}\right) B_t^{-1},$$

where A_i^{st} denotes the st-th element of A_i^{-1}.

Proof. Let the $N \times T$ matrix U be defined as in (1) with T columns $u_1, u_2, \dots, u_T$ and N rows $\tilde{u}_1', \tilde{u}_2', \dots, \tilde{u}_N'$. Further, let

$$\bar{A} = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & A_N \end{pmatrix}, \qquad \bar{B} = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & B_T \end{pmatrix}.$$

Finally, let K denote the $NT \times NT$ commutation matrix that transforms $\text{vec}(E)$ into $\text{vec}(E')$, so that $K\text{vec}(E) = \text{vec}(E')$; see Magnus and Neudecker (1988, Sect. 3.7) and Magnus (1988) for further details. Note that K is a permutation matrix, hence orthogonal: $K' = K^{-1}$. Then,

$$u = \text{vec}(U) = \bar{B}\text{vec}(E) = \bar{B}K'\text{vec}(E'),$$

so that

$$\Omega = \text{var}(u) = \bar{B}K'\text{var}(\text{vec}E')K\bar{B}' = \bar{B}K'\bar{A}K\bar{B}'.$$

This implies that

$$|\Omega| = \prod_{t=1}^{T} |B_t|^2 \prod_{i=1}^{N} |A_i|$$

and

$$\Omega^{-1} = (\bar{B}')^{-1}K'\bar{A}^{-1}K\bar{B}^{-1}.$$

To obtain explicit expressions for the blocks of Ω^{-1}, we let p_i be the ith column of I_N and q_t the tth column of I_T. Then we can write $K = \sum_i \sum_t (p_i q_t' \otimes q_t p_i')$; see Magnus (1988, Thm. 3.2). Hence,

$$\begin{aligned} K'\bar{A}^{-1}K &= \sum_{i,j,h=1}^{N} \sum_{s,t=1}^{T} (q_s p_i' \otimes p_i q_s')(p_h p_h' \otimes A_h^{-1})(p_j q_t' \otimes q_t p_j') \\ &= \sum_{i,j,h=1}^{N} \sum_{s,t=1}^{T} (q_s p_i' p_h p_h' p_j q_t') \otimes (p_i q_s' A_h^{-1} q_t p_j') \\ &= \sum_{i=1}^{N} \sum_{s,t=1}^{T} (q_s q_t') \otimes (p_i q_s' A_i^{-1} q_t p_i') = \sum_{s,t=1}^{T} (q_s q_t') \otimes \sum_{i=1}^{N} A_i^{st} p_i p_i' \\ &= \sum_{s,t=1}^{T} (q_s q_t') \otimes \text{diag}\left(A_1^{st}, A_2^{st}, \ldots, A_N^{st}\right), \end{aligned}$$

from which the blocks Ω^{st} follow directly. ∎

We notice that in Theorem 4 all three objectives mentioned in the Introduction have been realized. There is heteroskedasticity (through A_i), there is time-specific spatial correlation (through B_t, e.g., by specifying $\varepsilon_t = W\varepsilon_t + u_t$, so that $B_t = I_N - W$, in this case constant), and there is household-specific autocorrelation (also through A_i).

5. SOME REMARKS ON THREE ERROR COMPONENTS

One may wonder whether a useful extension from two error components to three error components is possible. It turns out that this is not the case, and we briefly investigate why this is so.

A three error component model would consist of three independent errors leading to a sum of three variance matrices. In the simplest case we write the error u_{it} as

$$u_{it} = \upsilon_i + \zeta_t + \epsilon_{it},$$

where the υ_i are independent and identically distributed with mean zero and variance σ_υ^2 and, similarly, the ζ_t and ϵ_{it} are independent and identically distributed with mean zero and variance σ_ζ^2 and σ_ϵ^2, respectively. If we order the errors as in (1) and let $J_T = \imath_T\imath_T'/T$ and $J_N = \imath_N\imath_N'/N$, then we see that $\Omega = \text{var}(\text{vec}U)$ is given by

$$\begin{aligned}\Omega &= \sigma_\upsilon^2(\imath_T\imath_T' \otimes I_N) + \sigma_\zeta^2(I_T \otimes \imath_N\imath_N') + \sigma_\epsilon^2(I_T \otimes I_N)\\ &= T\sigma_\upsilon^2(J_T \otimes I_N) + N\sigma_\zeta^2(I_T \otimes J_N) + \sigma_\epsilon^2(I_T \otimes I_N)\\ &= J_T \otimes B_1 + (I_T - J_T) \otimes B_2,\end{aligned}$$

where

$$B_1 = (T\sigma_\upsilon^2 + N\sigma_\zeta^2 + \sigma_\epsilon^2)J_N + (T\sigma_\upsilon^2 + \sigma_\epsilon^2)(I_N - J_N)$$

and

$$B_2 = (N\sigma_\zeta^2 + \sigma_\epsilon^2)J_N + \sigma_\epsilon^2(I_N - J_N).$$

We now easily obtain the determinant and inverse of Ω using Lemma 1. What this shows is that the three error component model is a restriction rather than an extension of the two error component model, because the three error component model puts structure on B_1 and B_2 whereas the two error component model only requires these matrices to be positive definite.

Even if we allow for a more general three error component structure by writing

$$\Omega = aa' \otimes B + A \otimes bb' + \tau^2 A \otimes B,$$

then the determinant and inverse can still be easily computed, but again there is no gain because

$$\Omega = aa' \otimes B + A \otimes (bb' + \tau^2 B)$$

is simply a specialized expression of the two error component model.

Of course there are three error component structures

$$\Omega = A_1 \otimes B_1 + A_2 \otimes B_2 + A_3 \otimes B_3$$

that are more general than a two error component structure. But the determinant and inverse of such structures are not easily computable in general. Based on the previous analysis we conjecture that the determinant and inverse of a three error component variance matrix can be written in terms of matrices of order T and N if and only if it can be written as a two error component variance matrix.

6. CONCLUSIONS

In this note we have presented several possible ways in which the $TN \times TN$ variance matrix of panel data models can be specified, allowing for maximum flexibility under the constraint that the determinant and inverse of the variance matrix can be calculated from matrices of orders $T \times T$ and $N \times N$ only. We conclude that much more generality is possible than is typically applied in panel data specifications.

REFERENCES

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

Anselin, L. & A.K. Bera (1998) Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah & D.E.A. Giles (eds.), *Handbook of Applied Economic Statistics*, pp. 237–289. Marcel Dekker.

Arellano, M. (2003) *Panel Data Econometrics*. Oxford University Press.

Baltagi, B.H. (2001) *Econometric Analysis of Panel Data*. Wiley.

Baltagi, B.H. & J.M. Griffin (1988) A generalized error component model with heteroskedastic disturbances. *International Economic Review* 39, 745–753.

Baltagi, B.H. & B. Raj (1992) A survey of recent theoretical developments in the econometrics of panel data. *Empirical Economics* 17, 85–109.

Baltagi, B.H., S.H. Song, B.C. Jung, & W. Koh (2007) Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics* 140, 5–51.

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

Driscoll, J.C. & A.C. Kraay (1998) Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* 80, 549–560.

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

Li, Q. & T. Stengos (1994) Adaptive estimation in the panel data error component model with heteroskedasticity of unknown form. *International Economic Review* 35, 981–1000.

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

Magnus, J.R. (1988) *Linear Structures.* Griffin's Statistical Monographs and Courses 42. Edward Arnold and Oxford University Press.

Magnus, J.R. & H. Neudecker (1988) *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley. [Second edition, 1999.]

Searle, S.R. & H.V. Henderson (1979) Dispersion matrices for variance components models. *Journal of the American Statistical Association* 74, 465–470.

Wansbeek, T. & A. Kapteyn (1982) A class of decompositions of the variance-covariance matrix of a generalized error components model. *Econometrica* 50, 713–724.