



## Volume 32, Issue 4

### On computing generalized least squares and maximum-likelihood estimates of error-components models with incomplete panels and correlated disturbances

Robert F. Phillips  
*George Washington University*

#### Abstract

Calculation of the inverse of the error variance-covariance matrix is required for both feasible generalized least squares and maximum-likelihood estimation of the regression parameters in the two-way error-components model. Since in many applications this matrix can be quite large, efficient computational methods for inverting the matrix are therefore needed. Incomplete panels complicate calculation of the inverse, and, to date, an efficient method for calculating the inverse has not been provided for the two-way error-components model in which the disturbances are correlated and the panel is incomplete. This note rectifies this shortcoming.

---

I am grateful to an anonymous referee for useful suggestions.

**Citation:** Robert F. Phillips, (2012) "On computing generalized least squares and maximum-likelihood estimates of error-components models with incomplete panels and correlated disturbances", *Economics Bulletin*, Vol. 32 No. 4 pp. 3017-3024.

**Contact:** Robert F. Phillips - [rphil@gwu.edu](mailto:rphil@gwu.edu).

**Submitted:** September 06, 2012. **Published:** October 30, 2012.

## 1. Introduction

Both feasible generalized least squares (GLS) and maximum-likelihood estimation of the regression parameters in the two-way error-components model rely on calculating the inverse of the error variance-covariance matrix. Since that matrix can be quite large, the practicality, therefore, of both feasible GLS and maximum-likelihood would be significantly diminished in the absence of efficient methods for calculating the inverse. Efficient computation of the inverse for the balanced panel data case has been known for some time (see, e.g., Fuller and Battese 1974; Nerlove 1971a; Wallace and Hussain 1969). However, for the unbalanced, or incomplete, panel data case, it was not until Wansbeek and Kapteyn's (1989) paper that an efficient method for computing the inverse was found.

Wansbeek and Kapteyn's (1989) solution relied on an analytical expression for the inverse that required inverting numerically only a  $T \times T$  matrix, where  $T$  is the total number of time periods observed. To obtain this solution, Wansbeek and Kapteyn (1989) arranged the observations so that the time index ran slowly while the cross section index ran fast; that is, cross sections observed in the first year were stacked on top of cross sections observed in the second year, and so on. This arrangement of panel observations has subsequently been adopted in the literature on the unbalanced two-way error-components model (see, e.g., Baltagi *et al.* 2002, and Phillips 2012).

However, arranging the observations as first suggested by Wansbeek and Kapteyn (1989) obscures the time-series structure of the panel. Obscuring this time-series structure is unimportant for the case considered by Wansbeek and Kapteyn (1989). That paper focused on the case in which the error components are all uncorrelated.<sup>1</sup> But in order to analyze more general time-series structures, it simplifies the analysis if the panel is arranged so that the time-series for one cross section is stacked on top of the time series for the next cross section, and so on.

This note shows that, if cross-sectional time series are stacked one on top of another, then it is still possible to invert the error variance-covariance matrix for an incomplete panel by inverting numerically matrices that are no larger than  $T \times T$ . The note thus provides a result analogous to the result Wansbeek and Kapteyn (1989) obtained, but the result provided here is for how panel data are usually analyzed, which allows for a unified treatment of the incomplete and complete panel cases. Moreover, an important advantage of this approach is that it also facilitates analyzing time-series models with correlated error components. To illustrate this point, I derive the main result in the note allowing the regression disturbances for each cross section to have an arbitrary variance-covariance matrix, and then I consider in detail the case where the time-series for cross sections do not all start and end at the same time and the disturbances are generated by a first-order autoregressive (AR(1)) process. The note concludes with an example which illustrates the potentially dramatic improvements in computing speed made possible by the main result in the note.

---

<sup>1</sup>Davis (2002) studied a more general error-components model than the model considered in Wansbeek and Kapteyn (1989). However, like Wansbeek and Kapteyn (1989), Davis (2002) assumed the error components are all uncorrelated, with each other and among themselves. Moreover, Baltagi *et al.* (2002) and Phillips (2012) also focused on the case of uncorrelated error components. On the other hand, for the balanced panel data case, Karlsson and Skoglund (2004) studied the two-way error-components model with correlated time effects.

## 2. The model

The model studied in this note is

$$y_{it} = \mu + \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{w}'_i\boldsymbol{\alpha} + u_{it} \quad (1)$$

with a composite regression error  $u_{it} = a_i + \tau_t + v_{it}$ . Here  $\mathbf{x}_{it}$  is a  $K_x \times 1$  matrix of regressors that vary with  $t$  and possibly  $i$ ;  $\mathbf{w}_i$  is a  $K_w \times 1$  vector of time-invariant regressors;  $\boldsymbol{\beta}$  and  $\boldsymbol{\alpha}$  are vectors of unobserved parameters;  $a_i$  and  $\tau_t$  are unobserved cross-sectional and time specific effects; and  $v_{it}$  will be referred to as a disturbance term.

Let  $T_i$  denote the number of observations on the  $i$ th cross section, which may be strictly less than  $T$ , the total number of time periods. The first observation on a given cross section may not correspond to the first year for which there is data for other cross sections, nor need the last observation for that cross section correspond to the last year. Moreover, for a given cross section, there may be missing years between two observations.

Next, let  $\mathbf{y}_i$  denote the  $T_i \times 1$  vector of observations on  $y_{it}$  ordered from smallest to largest  $t$ , and let  $\mathbf{X}_i$  be the corresponding  $T_i \times K$  matrix of observations on  $\mathbf{x}'_{it}$ . Define the  $T_i \times 1$  vector of disturbances,  $\mathbf{v}_i$ , analogously. Moreover, set  $\mathbf{y} = (\mathbf{y}'_1, \dots, \mathbf{y}'_N)'$ ,  $\mathbf{X} = (\mathbf{X}'_1, \dots, \mathbf{X}'_N)'$ , and  $\mathbf{v} = (\mathbf{v}'_1, \dots, \mathbf{v}'_N)'$ . Finally, let  $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_N)'$ ;  $\mathbf{a} = (a_1, \dots, a_N)'$  and  $\boldsymbol{\tau} = (\tau_1, \dots, \tau_T)'$ ;  $\boldsymbol{\iota}_n$  is an  $n \times 1$  vector of ones with  $n$  denoting the number of observations;  $\Delta_1 = \text{diag}(\boldsymbol{\iota}_{T_1}, \dots, \boldsymbol{\iota}_{T_N})$ ; and  $\Delta_2 = (\mathbf{D}'_1, \dots, \mathbf{D}'_N)'$ , where  $\mathbf{D}_i$  is a  $T_i \times T$  matrix constructed by deleting the rows of the  $T \times T$  identity matrix corresponding to those time periods not observed for the  $i$ th cross section. Using these definitions, the model can be written as

$$\mathbf{y} = \boldsymbol{\iota}_n\mu + \mathbf{X}\boldsymbol{\beta} + \Delta_1\mathbf{W}\boldsymbol{\alpha} + \mathbf{u},$$

where  $\mathbf{u} = \Delta_1\mathbf{a} + \Delta_2\boldsymbol{\tau} + \mathbf{v}$ .

Let  $\sigma^2\boldsymbol{\Omega} = E(\mathbf{u}\mathbf{u}'|\mathbf{X}, \mathbf{W})$ . It is assumed that  $\boldsymbol{\Omega} = \boldsymbol{\Sigma} + \phi_a\Delta_1\Delta'_1 + \phi_\tau\Delta_2\Delta'_2$ , where  $\phi_a = \sigma_a^2/\sigma^2$ , and  $\phi_\tau = \sigma_\tau^2/\sigma^2$ . Also assume that  $\mathbf{v}_i$  and  $\mathbf{v}_j$  are independent for  $i \neq j$ . Then  $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\Sigma}_{T_1}, \dots, \boldsymbol{\Sigma}_{T_N})$ , where  $\sigma^2\boldsymbol{\Sigma}_{T_i} = E(\mathbf{v}_i\mathbf{v}'_i|\mathbf{X}, \mathbf{W})$ .

## 3. A lemma

Wansbeek and Kapteyn (1989) showed how to calculate the inverse of the error variance-covariance matrix by inverting only a  $T \times T$  matrix numerically. They obtained their result by assuming the error components are all uncorrelated and by stacking the observations for the first year on top of the observations for the second year, and so on. Alternatively, Lemma 1 shows one can allow for correlated and heteroskedastic  $v_{it}$ s and arrange the sample observations as a time series (with possibly missing cases) for each cross section, and then the time series for the first cross section is stacked on top of the time series for the second, and so on.

*Lemma 1.* Let  $\mathbf{C} = \boldsymbol{\Sigma}^{-1} - \phi_a\boldsymbol{\Sigma}^{-1}\Delta_1\boldsymbol{\Upsilon}\Delta'_1\boldsymbol{\Sigma}^{-1}$ , where  $\boldsymbol{\Upsilon} = \text{diag}(\lambda_{T_1}, \dots, \lambda_{T_N})$ , and  $\lambda_{T_i} = (1 + \phi_a\boldsymbol{\iota}'_{T_i}\boldsymbol{\Sigma}^{-1}_{T_i}\boldsymbol{\iota}_{T_i})^{-1}$ . Then

$$\boldsymbol{\Omega}^{-1} = \mathbf{C} - \phi_\tau\mathbf{C}\Delta_2(\mathbf{I}_T + \phi_\tau\Delta'_2\mathbf{C}\Delta_2)^{-1}\Delta'_2\mathbf{C}. \quad (2)$$

The proof is provided in the appendix.

According to Lemma 1,  $\Omega^{-1}$  can be calculated by calculating the inverses  $\Sigma^{-1}$  and  $(\mathbf{I}_T + \phi_\tau \Delta_2' \mathbf{C} \Delta_2)^{-1}$ . The matrix  $\Sigma^{-1}$  is block-diagonal, with diagonal blocks  $\Sigma_{T_i}^{-1}$  ( $i = 1, \dots, N$ ), all of which are no larger than the  $T \times T$  matrix  $\mathbf{I}_T + \phi_\tau \Delta_2' \mathbf{C} \Delta_2$ .

Moreover, for several well-known time-series models there is a known  $T_i \times T_i$  matrix  $\mathbf{P}_{T_i}$  such that  $\mathbf{P}_{T_i}' \mathbf{P}_{T_i} = \Sigma_{T_i}^{-1}$ . The simplest case is when, for each cross section, the  $v_{it}$ s are generated by a stationary AR(1) process, with autocorrelation coefficient  $\rho$ , and there are no gaps in the time series. When there are no gaps in the time series, the time series for cross sections may differ in when they start and end, but there are no missing years between two consecutive observations. This restriction rules out some applications, but it still captures many, such as panels where individuals attrit, or panels where the data for some countries do not extend as far back in time as others, or rotating panels, and so on. When there are no gaps in the time series, the  $T_i \times T_i$  matrix  $\mathbf{P}_{T_i}$  is the Prais-Winsten transformation matrix

$$\mathbf{P}_{T_i} = \begin{bmatrix} \sqrt{1 - \rho^2} & 0 & \cdots & 0 \\ -\rho & 1 & & \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & -\rho & 1 \end{bmatrix}. \quad (3)$$

Now let  $\mathbf{P} = \text{diag}(\mathbf{P}_{T_1}, \dots, \mathbf{P}_{T_N})$ . Then, in this case,

$$\mathbf{C} = \mathbf{P}' (\mathbf{I}_n - \phi_a \mathbf{P} \Delta_1 \Upsilon \Delta_1' \mathbf{P}') \mathbf{P} \quad (4)$$

and

$$\lambda_{T_i} = (1 + \phi_a \iota_{T_i}' \mathbf{P}_{T_i}' \mathbf{P}_{T_i} \iota_{T_i})^{-1}.^2 \quad (5)$$

The AR(1) model without gaps in any of the time series is not the only case for which there is a known matrix  $\mathbf{P}_{T_i}$  satisfying  $\mathbf{P}_{T_i}' \mathbf{P}_{T_i} = \Sigma_{T_i}^{-1}$ . Baltagi and Wu (1999) provide the  $\mathbf{P}_{T_i}$  matrix for the AR(1) model with time series gaps. For the  $\mathbf{P}_{T_i}$  matrix for the MA(1) model, the AR(2) model, and the AR(4) model for quarterly data, all without gaps in the time series, see Balestra (1980), Lempers and Kloek (1973), and Thomas and Wallis (1971), respectively. For any of these cases, we can set  $\mathbf{P} = \text{diag}(\mathbf{P}_{T_1}, \dots, \mathbf{P}_{T_N})$ , and calculate  $\mathbf{C}$  and  $\lambda_{T_i}$  as in (4) and (5).

Lemma 1 increases computing speed by reducing the size of the matrices that are multiplied and inverted. To see this, let  $\mathbf{Z} = (\iota_n, \mathbf{X}, \Delta_1 \mathbf{W})$ , and consider that computing the GLS estimate  $(\mathbf{Z}' \Omega^{-1} \mathbf{Z})^{-1} \mathbf{Z}' \Omega^{-1} \mathbf{y}$  directly involves matrix multiplication and inversion of matrices that can, in some applications, be quite large. Lemma 1, on the other hand, allows us to replace the operations on these large matrices with sums, the terms of which involve operations on much smaller matrices.

To see this result, consider the evaluation of  $\mathbf{Z}' \Omega^{-1} \mathbf{Z}$ . From formulas (2), (4), (5), and straightforward calculations we get

$$\mathbf{Z}' \Omega^{-1} \mathbf{Z} = \mathbf{S}_{ZZ} - \phi_\tau \mathbf{S}_{Z\Delta_2} (\mathbf{I}_T + \phi_\tau \mathbf{S}_{\Delta_2\Delta_2})^{-1} \mathbf{S}_{\Delta_2Z}, \quad (6)$$

<sup>2</sup>For the AR(1) model,  $\lambda_{T_i}$  simplifies to  $\lambda_{T_i} = \{1 + \phi_a [1 - \rho^2 + (T_i - 1)(1 - \rho)^2]\}^{-1}$ .

where

$$\begin{aligned} \mathbf{S}_{ZZ} &= \sum_{i=1}^N \mathbf{Z}'_i \mathbf{P}'_{T_i} (\mathbf{I}_{T_i} - \phi_a \lambda_{T_i} \mathbf{P}_{T_i} \mathbf{J}_{T_i} \mathbf{P}'_{T_i}) \mathbf{P}_{T_i} \mathbf{Z}_i, \\ \mathbf{S}_{Z\Delta_2} &= \sum_{i=1}^N \mathbf{Z}'_i \mathbf{P}'_{T_i} (\mathbf{I}_{T_i} - \phi_a \lambda_{T_i} \mathbf{P}_{T_i} \mathbf{J}_{T_i} \mathbf{P}'_{T_i}) \mathbf{P}_{T_i} \mathbf{D}_i, \\ \mathbf{S}_{\Delta_2\Delta_2} &= \sum_{i=1}^N \mathbf{D}'_i \mathbf{P}'_{T_i} (\mathbf{I}_{T_i} - \phi_a \lambda_{T_i} \mathbf{P}_{T_i} \mathbf{J}_{T_i} \mathbf{P}'_{T_i}) \mathbf{P}_{T_i} \mathbf{D}_i, \end{aligned}$$

$\mathbf{J}_{T_i} = \iota_{T_i} \iota'_{T_i}$ , and  $\mathbf{S}_{\Delta_2 Z} = \mathbf{S}'_{Z\Delta_2}$ . Similarly,

$$\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{y} = \mathbf{S}_{Zy} - \phi_\tau \mathbf{S}_{Z\Delta_2} (\mathbf{I}_T + \phi_\tau \mathbf{S}_{\Delta_2\Delta_2})^{-1} \mathbf{S}_{\Delta_2 y}, \quad (7)$$

where

$$\begin{aligned} \mathbf{S}_{Zy} &= \sum_{i=1}^N \mathbf{Z}'_i \mathbf{P}'_{T_i} (\mathbf{I}_{T_i} - \phi_a \lambda_{T_i} \mathbf{P}_{T_i} \mathbf{J}_{T_i} \mathbf{P}'_{T_i}) \mathbf{P}_{T_i} \mathbf{y}_i, \\ \mathbf{S}_{\Delta_2 y} &= \sum_{i=1}^N \mathbf{D}'_i \mathbf{P}'_{T_i} (\mathbf{I}_{T_i} - \phi_a \lambda_{T_i} \mathbf{P}_{T_i} \mathbf{J}_{T_i} \mathbf{P}'_{T_i}) \mathbf{P}_{T_i} \mathbf{y}_i. \end{aligned}$$

To get a sense of how much the formula on the right-hand side of (6) speeds up computations, note that, although the sums  $\mathbf{S}_{ZZ}$ ,  $\mathbf{S}_{Z\Delta_2}$ , and  $\mathbf{S}_{\Delta_2\Delta_2}$  each have  $N$  terms, the dimensions of the matrices in these sums do not depend on  $N$ , nor does the dimension of  $\mathbf{I}_T + \phi_\tau \mathbf{S}_{\Delta_2\Delta_2}$ . Thus, for a given  $T$ , if we use the formula on the right-hand side of (6) to calculate  $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z}$ , the number of arithmetical operations increases with  $N$  at the same rate  $N$  increases. On the other hand, the number of arithmetical operations required just to invert  $\boldsymbol{\Omega}$  increases with  $N$  at a much faster rate than the rate at which  $N$  increases.<sup>3</sup> This is easiest to see in the balanced panel case, that is, when  $T_i = T$  ( $i = 1, \dots, N$ ). In this case, for given  $T$ , the number of arithmetical operations required to invert  $\boldsymbol{\Omega}$  increases with  $N$  at a rate on the order of  $N^\delta$ , with  $\delta$  no smaller than two.<sup>4</sup>

#### 4. An example

In order to illustrate the gains in computing speed afforded by exploiting Lemma 1, I generated samples of explanatory values and calculated  $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z}$  both by exploiting the formula on the right-hand side of (6) and by calculating  $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z}$  directly. Since the focus here is on computation rather than estimation, the values of  $\phi_a$  and  $\phi_\tau$  were taken to be known and were both set to one. For  $\mathbf{P}_{T_i}$ , the Prais-Winsten transformation matrix in (3) was used with  $\rho = 0.5$ .

<sup>3</sup>Similar observations apply for computing  $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{y}$ .

<sup>4</sup>This follows from the fact that the number of arithmetical operations required to invert  $\boldsymbol{\Omega}$  is of order  $n^\delta$ . Thus, if  $T_i = T$ , for all  $i$ , and consequently  $n^\delta = N^\delta T^\delta$ , the number of arithmetical operations required to invert  $\boldsymbol{\Omega}$  increases with  $N$  at a rate on the order of  $N^\delta$ . As for the size of  $\delta$ , the best lower bound is two (see Pan 1984, p. 1).

Table 1: Ratios of computing times (fast over slow)

<i>N</i>	<i>T</i> = 10				<i>T</i> = 50			
	10	50	100	500	10	50	100	500
Ratio	0.2738	0.0348	0.0090	$4.1 \times 10^{-4}$	0.1402	0.0099	0.0026	$9.4 \times 10^{-5}$

Regressor values were generated using a model similar to that used by Nerlove (1971b). Specifically, I set

$$x_{it} = 0.1t + 0.5x_{i,t-1} + \omega_{it}, \quad (t = 1, \dots, T_i, i = 1, \dots, N),$$

where the  $\omega_{it}$ s were generated independently as uniform variates over the interval  $(-0.5, 0.5)$ , and  $x_{i0} = 5 + 10\omega_{i0}$ . The  $t$ th row of  $\mathbf{Z}_i$  was set equal to  $(1, x_{it})$  ( $t = 1, \dots, T_i, i = 1, \dots, N$ ).

Several different  $(T, N)$  combinations were considered, where  $T$  is the maximum number of observed time periods and  $N$  is the total number of cross sections. For the  $(T, N)$  combinations, I set  $N$  equal to 10, 50, 100, and 500, for  $T = 10$  and  $T = 50$ . Moreover, the panels were unbalanced. Specifically, for each sample,  $T_i$  was set as  $T_i = T/2$ , for  $i = 1, \dots, N/2$ , and  $T_i = T$ , for  $i = N/2 + 1, \dots, N$ . Hence, for the smallest and largest samples, the total number of observations,  $n$ , was 75 and 18,750.

For each sample, the time required to calculate  $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z}$  with and without taking advantage of Lemma 1 was recorded. Table 1 gives the ratio of the time required to calculate  $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z}$  using the formula on the right-hand side of (6) over the time required to calculate  $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z}$  by computing the inverse of  $\boldsymbol{\Omega}$ , postmultiplying  $\boldsymbol{\Omega}^{-1}$  by  $\mathbf{Z}$ , and premultiplying by  $\mathbf{Z}'$ .<sup>5</sup>

Consistent with the observations at the end of Section 3, the benefits of applying Lemma 1 to calculate  $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z}$  depended on the size of the matrices involved, with the relative speed increasing with the size of the matrices. Indeed, for the case in which the matrices  $\boldsymbol{\Omega}$  and  $\mathbf{Z}$  were largest ( $T = 50, N = 500$ ), applying Lemma 1 produced an algorithm that was over 10,600 times faster than an algorithm consisting of simply inverting  $\boldsymbol{\Omega}$ , postmultiplying by  $\mathbf{Z}$ , and premultiplying by  $\mathbf{Z}'$ .

### Appendix: Proof of Lemma 1

Eq. (2) follows from the identity

$$(\mathbf{A} + \mathbf{UBV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{I} + \mathbf{BVA}^{-1}\mathbf{U})^{-1}\mathbf{BVA}^{-1} \quad (8)$$

<sup>5</sup>All computations were performed using GAUSS on a MacBook Pro. Computing time was estimated elapsed time. The GAUSS command “hsec” was used to calculate the number of hundredths of a second since midnight before and after calculations were executed. The difference between these two times was the estimated elapsed time.

(see Henderson and Searle 1981). Set  $\mathbf{A} = \Sigma + \phi_a \Delta_1 \Delta_1' + \phi_\tau \Delta_2 \Delta_2'$ ,  $\mathbf{U} = \mathbf{V}' = \Delta_2$ , and  $\mathbf{B} = \phi_\tau$ . Then, according to Eq. (8),

$$\begin{aligned}\Omega^{-1} &= (\Sigma + \phi_a \Delta_1 \Delta_1' + \phi_\tau \Delta_2 \Delta_2')^{-1} \\ &= \mathbf{A}^{-1} - \phi_\tau \mathbf{A}^{-1} \Delta_2 (\mathbf{I}_T + \phi_\tau \Delta_2' \mathbf{A}^{-1} \Delta_2)^{-1} \Delta_2' \mathbf{A}^{-1}.\end{aligned}$$

Moreover, we can evaluate  $\mathbf{A}^{-1}$  using (8) again:

$$\mathbf{A}^{-1} = \Sigma^{-1} - \phi_a \Sigma^{-1} \Delta_1 (\mathbf{I}_N + \phi_a \Delta_1' \Sigma^{-1} \Delta_1)^{-1} \Delta_1' \Sigma^{-1}.$$

Finally, straightforward calculations give that  $(\mathbf{I}_N + \phi_a \Delta_1' \Sigma^{-1} \Delta_1)^{-1} = \Upsilon$ . Collecting the proceeding observations establishes Eq. (2).

### References

- Balestra, P. (1980) "A Note on the Exact Transformation Associated with the First-Order Moving Average Process" *Journal of Econometrics* 14, 381-394.
- Baltagi, B. H., Song, S. H., and B. C. Jung (2002) "A Comparative Study of Alternative Estimators for the Unbalanced Two-Way Error Component Regression Model" *Econometrics Journal* 5, 480-493.
- Baltagi, B. H. and P. X. Wu (1999) "Unequally Spaced Panel Data Regressions with AR(1) Disturbances" *Econometric Theory* 15, 814-823.
- Davis, P. (2002) "Estimating Multi-Way Error Components Models with Unbalanced Data Structures" *Journal of Econometrics* 106, 67-95.
- Fuller, W. A. and G. E. Battese (1974) "Estimation of Linear Models with Crossed-Error Structure" *Journal of Econometrics* 2, 67-78.
- Henderson, H. V. and S. R. Searle (1981) "On Deriving the Inverse of a Sum of Matrices" *SIAM Review* 23, 53-60.
- Karlsson, S. and J. Skoglund (2004) "Maximum-Likelihood Based Inference in the Two-Way Random Effects Model with Serially Correlated Time Effects" *Empirical Economics* 29, 79-88.
- Lempers, F. B. and T. Kloek (1973) "On a Simple Transformation for Second-Order Autocorrelated Disturbances in Regression Analysis" *Statistica Neerlandica* 27, 69-75.
- Nerlove, M. (1971a) "A Note on Error Components Models" *Econometrica* 39, 383-396.
- Nerlove, M. (1971b) "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections" *Econometrica* 39, 359-382.

Pan, V. (1984) *How to Multiply Matrices Faster*, Springer-Verlag: New York.

Phillips, R. F. (2012) “On Computing Maximum-Likelihood Estimates of the Unbalanced Two-Way Random-Effects Model” *Communications in Statistics – Simulation and Computation* 41, 1921-1927.

Thomas, J. J. and K. F. Wallis (1971) “Seasonal Variation in Regression Analysis” *Journal of the Royal Statistical Society, Series A (General)* 134, 57-72.

Wallace, T. D. and A. Hussain (1969) “The Use of Error Components Models in Combining Cross Section with Time Series Data” *Econometrica* 37, 55-72.

Wansbeek, T. and A. Kapteyn (1989) “Estimation of the Error-Components Model with Incomplete Panels” *Journal of Econometrics* 41, 341-361.