Tilburg University

Estimation of the error-components model with incomplete panels

Wansbeek, T.J.; Kapteyn, A.J.

Publication date:
1990

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Wansbeek, T. J., & Kapteyn, A. J. (1990). *Estimation of the error-components model with incomplete panels*. (Reprint series / CentER for Economic Research; Vol. 18). Unknown Publisher.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

CBM
entER
for
omic Research

8823
1990

18

REPRINT

Estimation of the Error-
Components Model with
Incomplete Panels

by
Tom Wansbeek and
Arie Kapteyn

Reprinted from Journal of Econometrics,
Vol. 41, No. 3, 1989

Reprint Series
no. 18

CENTER FOR ECONOMIC RESEARCH

Scientific Council

Eduard Bomhoff	Erasmus University Rotterdam
Willem Buiter	Yale University
Jacques Drèze	Université Catholique de Louvain
Theo van de Klundert	Tilburg University
Simon Kuipers	Groningen University
Jean-Jacques Laffont	Université des Sciences Sociales de Toulouse
Merton Miller	University of Chicago
Stephen Nickell	University of Oxford
Pieter Ruys	Tilburg University
Jacques Sijben	Tilburg University

Board

Anton P. Barten, director
Eric van Damme
John Driffill
Arie Kapteijn
Frederick van der Ploeg

Research Staff and Resident Fellows

Anton P. Barten
Eric van Damme
John Driffill
Arie Kapteyn
Hugo Keuzenkamp
Pieter Kop Jansen
Jan Magnus
Frederick van der Ploeg

Address: Hogeschoollaan 225, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
Phone : +3113663050
Telex : 52426 kub nl
Telefax: +3113663066
E-mail : "CENTER@HTIKUB5.BITNET"

Center

for

Economic Research

REPRINT

Estimation of the Error-
Components Model with
Incomplete Panels

by

Tom Wansbeek and
Arie Kapteyn

Reprinted from Journal of Econometrics,
Vol. 41, No. 3, 1989

Reprint Series
no. 18

ESTIMATION OF THE ERROR-COMPONENTS MODEL WITH INCOMPLETE PANELS*

Tom WANSBEEK

Groningen University, Groningen, The Netherlands

Arie KAPTEYN

Tilburg University, Tilburg, The Netherlands

Received March 1986, final version received September 1988

The error-components model (ECM) is probably the most frequently used approach to analyze panel data in econometrics. When the panel is incomplete, which is the rule rather than the exception when the data come from large-scale surveys, standard estimation methods cannot be applied. We first discuss estimation in the fixed-effects analogue of the ECM, and then present two estimators (quadratic unbiased and maximum likelihood) for the ECM. Some simulation results are given to assess finite-sample properties and computational burden of the various methods.

1. Introduction

Analysis of panel data (i.e., time series of cross-sections) by means of the error-components model (ECM) has attracted a lot of attention in econometrics; see, e.g., Balestra and Nerlove (1966), Wallace and Hussain (1969), Nerlove (1971a, b), Mazodier (1972), Fuller and Battese (1974), Taylor (1977), Mazodier (1978), Baltagi (1981), Wansbeek and Kapteyn (1982), and Dielman (1983). The recent monograph by Hsiao (1986) presents an excellent overview. A problem that often occurs in practice, but which is by and large ignored in the literature, is the phenomenon of missing observations, i.e., not all cross-sectional units are observed during all time periods. If that occurs, we have what we will call an *incomplete panel* or *incomplete data*. Note that we use the word 'incomplete' to denote the absence of *all* information for a certain cross-section unit for a certain time period, and not to be the absence of information on some variables only.

*Detailed and helpful comments from two anonymous referees are gratefully acknowledged. We thank Anton Markink and Paul Flapper for their expert programming support, Brent Moulton and Theo Nijman for their comments on an earlier version, and Jaap Verhees for checking the many algebraic manipulations.

In their pioneering study, Balestra and Nerlove (1966) analyzed panel data where the cross-section units were states of the USA, the time periods being years. The data set was constructed by the authors from existing sources and in such a context incompleteness in the above sense will not be a major concern. However, with the growing attention for micro-data in econometrics, data are increasingly obtained from large-scale surveys and there incompleteness will be the rule rather than the exception. For example, individuals may disappear from a panel after a few waves because they refuse to cooperate any longer, because they leave the household that is participating in the panel, or by death. Moreover, incompleteness may in fact be part of a sample design in which part of a panel is replaced by new units in each wave (rotating panels). The advantage of such a design is that it eases the task of respondents, thereby reducing attrition from the panel. Since attrition generally causes selection bias [see, e.g., Hausman and Wise (1977)], which can only be remedied by means of fairly elaborate and computationally costly models, a sample design which minimizes attrition is of practical interest.

The absence of some observations makes most of the results obtained in the error-components literature inapplicable. A solution was suggested by Fuller and Battese (1974), in which a dummy exogenous variable is added to the set of regressors for each missing observation. After this amendment of the data set, the usual methods can be applied. In many practical cases, however, this means that hundreds of regressors have to be added. This is clearly computationally impractical and an alternative is called for.

Biörn (1981) seems to be the first to discuss error-components models with missing observations. He discusses maximum-likelihood (ML) estimation in the case where a fixed proportion of the sample is replaced with each new wave, and does not allow for observations that may be missing randomly. He writes down the likelihood of this model, but then observes that it involves both the inverse and the determinant of the error-covariance matrix. Given the size of this matrix in practice, ML estimation without further provisions does not seem to be of practical value. Hence, he concentrates on the special case where at each wave exactly half of the sample is replaced and where there is no time-specific effect. Then, the inverse and the determinant can be worked out rather easily. Baltagi (1985) studies some aspects of the error-covariance matrix for a general 'missing pattern' in a model without time-specific effects.

In this paper we will first consider the relatively simple case where the effects (pertaining to the cross-section units and to the time periods) are fixed rather than random, as in the ECM. This case is dealt with in section 2. These results are of interest in themselves and are used, after section 3 where some basic aspects of the random-effects model are discussed, in section 4 as a starting point for deriving quadratic unbiased estimators in the (random-effects) ECM. Section 5 discusses ML estimation of the ECM. Section 6 contains some simulation results, and section 7 concludes. Some of the more

technical points are relegated to appendices; the subject of incomplete panels is closely related to the topic of 'unbalanced data' in the ANOVA literature, from which it is well-known that the results are invariably rather messy.

For clarity, it must be emphasized that some interesting problems lie outside the scope of the present paper. We only consider the single-equation model, and do not look at SUR-type or simultaneous equations. The only role that 'time' plays in the ECM is via the (supposedly independent) time-specific effects and we do not look at lagged endogenous variables or a more sophisticated modelling of the time component in the error structure; nor do we go into the problems caused by selective nonresponse. For all these extensions, we first need the basic apparatus given in the present paper.

2. The fixed-effects model

Consider the following regression model:

$$y_{ht} = \alpha_h + \gamma_t + x'_{ht}\beta + u_{ht}, \quad (2.1)$$

where h ($h = 1, \dots, H$) denotes, say, households, and t ($t = 1, \dots, T$), say, years; x_{ht} is a k -vector of explanatory variables; and β a k -vector of parameters. The error term u_{ht} has the usual 'classical' properties: its variance is σ^2 ; α_h and γ_t are fixed constants. Therefore, we will sometimes refer to this model as the 'fixed-effects' (FE) model. When we have a *complete* panel, i.e., all households h are observed for all years t , it is well-known that OLS in (2.1) is equivalent to OLS in the transformed model

$$y_{ht} - y_{h.} - y_{.t} + y_{..} = (x_{ht} - x_{h.} - x_{.t} + x_{..})'\beta + e_{ht}, \quad (2.2)$$

where a dot in the place of an index denotes the average over that index. A way to prove this result is to first write (2.1) in vector format,

$$y = (\iota_T \otimes I_H)\alpha + (I_T \otimes \iota_H)\gamma + X\beta + u. \quad (2.3)$$

Note that we have ordered the observations such that the data on the H households are ordered in T consecutive sets; the index t 'runs slowly' and the index h 'runs fast'. In (2.3), it is simple to show that the projector perpendicular to the regressors corresponding to the household and time effects, viz. $(\iota_T \otimes I_H, I_T \otimes \iota_H)$, is given by $E_T \otimes E_H$, where $E_T \equiv I_T - \iota_T \iota_T' / T$ and E_H is defined analogously. Moreover, it is rather straightforward to see that application of this projector to both sides of (2.1) effectuates the transformation shown in (2.2).

When we have incomplete data, these simple projection and transformation results no longer hold. We will now derive the corresponding more general results for the incomplete case.

Let N_t ($N_t \leq H$) be the number of observed households in year t . Let $N \equiv \sum_t N_t$. Let D_t be the $(N_t \times H)$ matrix obtained from the $(H \times H)$ identity matrix from which rows corresponding to households *not* observed in year t have been omitted, and consider

$$Z \equiv \begin{pmatrix} Z_1 & Z_2 \\ N \times H & N \times T \end{pmatrix} \equiv \begin{bmatrix} D_1 & D_1 \iota_H & & \\ \vdots & & \ddots & \\ D_T & & & D_T \iota_H \end{bmatrix}. \quad (2.4)$$

The matrix Z gives the dummy-variable structure for the incomplete-data model. (For complete data, $Z_1 = \iota_T \otimes I_H$, $Z_2 = I_T \otimes \iota_H$.)

Next, let

$$\Delta_H \equiv Z_1' Z_1, \quad \Delta_T \equiv Z_2' Z_2, \quad A \equiv Z_2' Z_1, \quad (2.5)$$

where Δ_H is the diagonal $(H \times H)$ matrix with h th element indicating the number of years for which the h th household has been observed; Δ_T is the diagonal $(T \times T)$ matrix with t th element the number of observations in year t ; and A is the $(T \times H)$ matrix of zeros and ones indicating the absence or presence of a household in a certain year. (For complete data, $\Delta_H = T I_H$, $\Delta_T = H I_T$, and $A = \iota_T \iota_H'$.) Further define

$$\bar{Z} \equiv Z_2 - Z_1 \Delta_H^{-1} A' \quad \left[= (I_N - Z_1 (Z_1' Z_1)^{-1} Z_1') Z_2 \right], \quad (2.6)$$

$$Q \equiv \Delta_T - A \Delta_H^{-1} A' \quad \left[= Z_2' \bar{Z} \right]. \quad (2.7)$$

In the complete data case, $Q = H E_T$. In the incomplete data case, Q has no specific structure. If each household is observed at least twice (and $H \geq T$), $\text{rank}(Q) = T - 1$. In order to avoid unnecessary complications, we assume in this section that this condition holds. (If a certain household is observed only once, it conveys no useful information as we then have the case of a single observation with its own dummy variable.)

Now consider the following matrix:

$$P \equiv P_1 - P_2 \equiv (I_N - Z_1 \Delta_H^{-1} Z_1') - \bar{Z} Q^{-1} \bar{Z}'. \quad (2.8)$$

Lemma. P is the projection matrix onto the null-space of Z .

Proof. The proof is in three steps:

- (i) P is idempotent: $Z_1' \bar{Z} = 0$ [clear from (2.6)], $\bar{Z}' \bar{Z} = Z_2' \bar{Z} = Q$. As $\bar{Z} Q^{-1} Q = \bar{Z}$, $P_2^2 = \bar{Z} Q^{-1} Q Q^{-1} \bar{Z}' = \bar{Z} Q^{-1} \bar{Z}' = P_2$. Also, $P_1^2 = P_1$ and $P_1 P_2 = P_2 P_1 = P_2$. So $P^2 = P_1^2 + P_2^2 - P_1 P_2 - P_2 P_1 = P_1 - P_2 = P$.

- (ii) $PZ = 0$: $PZ = P_1(Z_1, Z_2) - P_2(Z_1, Z_2) = (0, \bar{Z}) - (0, \bar{Z}Q^{-1}Q) = 0$.
- (iii) $\text{rank}(P) + \text{rank}(Z) = N$: $\text{rank}(P) = \text{tr}(P) = \text{tr}(P_1) - \text{tr}(P_2) = (N - H) - \text{tr}(Q^{-1}\bar{Z}'\bar{Z}) = (N - H) - \text{tr}(Q^{-1}Q) = (N - H) - (T - 1)$; $\text{rank}(Z) = H + T - 1$.

Together, (i), (ii), and (iii) prove the lemma [e.g., Balestra (1973, lemma 9)].
Q.E.D.

So, P generalizes the expression $E_T \otimes E_H$ to the incomplete-data model. Note that there is an asymmetry in the way that we deal with both dimensions (households and years): P contains the generalized inverse of the $(T \times T)$ matrix Q for which no closed-form expression is available in general. Alternatively, we could have derived an expression for P that contains the inverse of an analogous $(H \times H)$ matrix, but as $H \gg T$ in most practical situations, our choice is the most favorable one from the point of view of computation. The asymmetry and its inherent lack of elegance is the kind of uncomeliness that one has to face when dealing with incomplete or unbalanced data.

How to use P ? We will give the generalization of the transformation given in (2.2). Let $v(N \times 1)$ denote a vector of variables occurring in the regression equation; in (2.3), $v = y$ or v is a column of X . We are interested in the form of Pv . Let

$$\phi_1 \equiv Z_1'v \quad (H \times 1) \quad (2.9)$$

and

$$\phi_2 \equiv Z_2'v \quad (T \times 1) \quad (2.10)$$

denote the sum of elements of v over years and households, respectively, and let

$$\bar{\phi} \equiv Q^{-1}\bar{Z}v = Q^{-1}(\phi_2 - A\Delta_H^{-1}\phi_1). \quad (2.11)$$

(The choice of generalized inverse is arbitrary.) Now

$$Pv = v - Z_1\Delta_H^{-1}Z_1'v - \bar{Z}Q^{-1}\bar{Z}'v = v - Z_1\Delta_H^{-1}\phi_1 - \bar{Z}\bar{\phi}. \quad (2.12)$$

In scalar format this reads as

$$(Pv)_{ih} = v_{ih} - \frac{1}{\delta_h}\phi_{1h} + \frac{1}{\delta_h}a_h'\bar{\phi} - \bar{\phi}_i, \quad (2.13)$$

with a_h the h th column of A and δ_h the h th diagonal element of Δ_H . So, OLS on data from an incomplete panel with fixed year and household effects

amounts to OLS in a model without these effects when the variables have been transformed according to (2.13). Then using a standard regression package for the transformed data one should not forget to adjust the standard errors and the R^2 printed by the program for the loss of degrees of freedom. When there are k 'true' regressors (apart from the dummies), the printed standard errors should be multiplied by $\{(N-k)/(N-H-T+1-k)\}^{1/2}$. Analogously, the OLS residual variance estimate printed by the program should be multiplied by $(N-k)/(N-H-T+1-k)$ to obtain an unbiased estimate of σ^2 .

3. The random-effects model

In the usual ECM formulation the α_h and γ_t in (2.1) are i.i.d. random variables with mean zero, mutually independent and independent of the x_{ht} and u_{ht} . Let σ^2 be the variance of u_{ht} , then the covariance matrix of the composite error term $\varepsilon_{ht} \equiv u_{ht} + \alpha_h + \gamma_t$ is

$$\Omega = \sigma^2 I_N + \sigma_1^2 Z_1 Z_1' + \sigma_2^2 Z_2 Z_2', \quad (3.1)$$

with

$$\sigma_1^2 \equiv \text{var}(\alpha_h), \quad \sigma_2^2 \equiv \text{var}(\gamma_t).$$

For any sort of efficient estimation of the parameter vector β we need an expression for the inverse of Ω . This is given by the following:

Lemma.

$$\sigma^2 \Omega^{-1} = V - V Z_2 \tilde{Q}^{-1} Z_2' V, \quad (3.2)$$

where

$$V \equiv I_N - Z_1 \tilde{\Delta}_H^{-1} Z_1' \quad (N \times N), \quad (3.3)$$

$$\tilde{Q} \equiv \tilde{\Delta}_T - A \tilde{\Delta}_H^{-1} A' \quad (T \times T), \quad (3.4)$$

$$\tilde{\Delta}_H \equiv \Delta_H + \frac{\sigma^2}{\sigma_1^2} I_H \quad (H \times H), \quad (3.5)$$

$$\tilde{\Delta}_T \equiv \Delta_T + \frac{\sigma^2}{\sigma_2^2} I_T \quad (T \times T). \quad (3.6)$$

Proof.

$$Z_2'VZ_2 = \Delta_T - A\tilde{\Delta}_H^{-1}A' = \tilde{Q} - \frac{\sigma^2}{\sigma_2^2}I_T, \quad (3.7)$$

so

$$\tilde{Q} - Z_2'VZ_2 = \frac{\sigma^2}{\sigma_2^2}I_T. \quad (3.8)$$

From (3.3),

$$\begin{aligned} V^{-1} &= I_N + Z_1(\tilde{\Delta}_H - Z_1'Z_1)^{-1}Z_1' \\ &= I_N + Z_1(\tilde{\Delta}_H - \Delta_H)^{-1}Z_1' \\ &= I_N + \frac{\sigma_1^2}{\sigma^2}Z_1Z_1', \end{aligned} \quad (3.9)$$

so inverting the expression in (3.2) yields

$$\begin{aligned} (V - VZ_2\tilde{Q}^{-1}Z_2'V)^{-1} &= V^{-1} + Z_2(\tilde{Q} - Z_2'VZ_2)^{-1}Z_2' \\ &= I_N + \frac{\sigma_1^2}{\sigma^2}Z_1Z_1' + \frac{\sigma_2^2}{\sigma^2}Z_2Z_2' \\ &= \sigma^{-2}\Omega. \quad \text{Q.E.D.} \end{aligned} \quad (3.10)$$

The expression for Ω^{-1} is somewhat messy and is asymmetric in households and years. In contrast with the complete-data case, no closed-form expression for Ω^{-1} (or for the eigenvalues and eigenvectors) can be given in general. This is only possible for some very specific and 'neat' patterns of missing data. However, for practical purposes the expression for Ω^{-1} is quite useful as compared to the situation where Ω is inverted numerically. The aspect of main interest is the *computational complexity*, i.e., the number of computations, and hence the computing time, as a function of the number of observations. In our case, where T typically is a small number and N and H are large, we are interested in the computing time of the GLS estimator of the regression coefficients in terms of N and H , and take T to be a constant.

We can make the following observations. The computation of $\hat{\beta}$ involves a constant number of evaluations of the type $f'\Omega^{-1}g$, with f and g being

N -vectors (f and g may be one of the various regressors or the regressand). This, in its turn, involves a constant number of evaluations of the type $f'Vg$ [cf. (3.2)], where f and g now also comprise the columns of Z_2 . Now, $f'Vg = f'g - f'Z_1\bar{\Delta}_H^{-1}Z_1'g$; $f'g$ can be computed in $O(N)$ time and $f'Z_1\bar{\Delta}_H^{-1}Z_1'g$ in $O(H)$ time [at least, when the elements of f and g are displayed in a $(T \times H)$ matrix, then $f'Z_1$ and $Z_1'g$ are simply the H -vector of column totals, computable in $O(H)$ time]. If the structure of Ω would not have been exploited, Ω would have to be inverted numerically, which requires in principle $O(H^3)$ time. (The latter statement neglects recent developments in complexity theory, which allow for a reduction of the exponent 3 to a somewhat lower figure. Up until now, this development has theoretical significance only.)

The implication of the above is that the expressions used in inverting Ω may look a little deterring, but that they enable efficient computation (viz. computing time linear in H) of the GLS estimator of the regression coefficients.

4. Quadratic estimators of the variance components

Statistically efficient estimation of the ECM can take place along two lines. One is to use ML (see section 5). The other is to estimate the variance components (σ^2 , σ_1^2 , and σ_2^2) by a quadratic unbiased estimation method (QUE; below the E in QUE can also stand for estimator) and to estimate the regression coefficients by GLS with these estimates inserted in Ω . In this section we derive QUE's for σ_1^2 and σ_2^2 . The estimator for σ^2 from the FE model is unbiased under RE assumptions as well, so we concentrate on σ_1^2 and σ_2^2 .

An intuitively appealing approach to derive QUE's for σ_1^2 and σ_2^2 is to estimate the (FE) model and to use the FE residuals, averaged over households or averaged over years, as the basic ingredients. By residuals we mean in the present context the residuals with respect to the X -part of the FE model, not those with respect to all regressors (X and Z); the latter are not informative about σ_1^2 and σ_2^2 as the variation of interest is projected out.

We first assume that X does not contain a vector of ones. Let e be the N -vector of FE residuals, i.e., $e \equiv y - Xb$ with b the FE estimate of β , and let

$$q_H \equiv e'Z_2\Delta_T^{-1}Z_2'e, \quad (4.1)$$

$$q_T \equiv e'Z_1\Delta_H^{-1}Z_1'e, \quad (4.2)$$

$$k_H \equiv \text{tr}(X'PX)^{-1}X'Z_2\Delta_T^{-1}Z_2'X, \quad (4.3)$$

$$k_T \equiv \text{tr}(X'PX)^{-1}X'Z_1\Delta_H^{-1}Z_1'X. \quad (4.4)$$

Lemma.

$$E(q_H) = (T + k_H)\sigma^2 + T\sigma_1^2 + N\sigma_2^2, \quad (4.5)$$

$$E(q_T) = (H + k_T)\sigma^2 + N\sigma_1^2 + H\sigma_2^2. \quad (4.6)$$

Proof. We first prove (4.5). Let

$$M \equiv I_N - X(X'PX)^{-1}X'P, \quad (4.7)$$

then by definition $e = My = M\varepsilon$. As $PZ_1 = 0$ and $PZ_2 = 0$, $P\Omega = \sigma^2P$, so

$$M\Omega = \Omega - \sigma^2X(X'PX)^{-1}X'P, \quad (4.8)$$

$$\begin{aligned} M\Omega M' &= \Omega - \sigma^2X(X'PX)^{-1}X'P - \sigma^2PX(X'PX)^{-1}X' \\ &\quad + \sigma^2X(X'PX)^{-1}X', \end{aligned} \quad (4.9)$$

$$\begin{aligned} Z_2'M\Omega M'Z_2 &= Z_2'\Omega Z_2 + \sigma^2Z_2'X(X'PX)^{-1}X'Z_2 \\ &= \sigma^2\Delta_T + \sigma_1^2AA' + \sigma_2^2\Delta_T + \sigma^2Z_2'X(X'PX)^{-1}X'Z_2. \end{aligned} \quad (4.10)$$

Using (4.10) and $\text{tr}(\Delta_T^{-1}AA') = T$,

$$\begin{aligned} E(q_H) &= E(e'Z_2\Delta_T^{-1}Z_2'e) \\ &= \text{tr} E(\Delta_T^{-1}Z_2'M\varepsilon\varepsilon'M'Z_2) \\ &= \text{tr}(\Delta_T^{-1}Z_2'M\Omega M'Z_2) \\ &= \text{tr}(\sigma^2I_T + \sigma_1^2\Delta_T^{-1}AA' + \sigma_2^2\Delta_T + \sigma^2\Delta_T^{-1}Z_2'X(X'PX)^{-1}X'Z_2) \\ &= (T + k_H)\sigma^2 + T\sigma_1^2 + N\sigma_2^2. \end{aligned} \quad (4.11)$$

The proof of (4.6) is analogous. Q.E.D.

We obtain QUE's for σ_1^2 and σ_2^2 when solving (4.5) and (4.6) for σ_1^2 and σ_2^2 and using an unbiased estimator for σ^2 . The latter may be based on

$$E(e'Pe) = \sigma^2(N - T - H + 1 - k), \quad (4.12)$$

i.e., the estimator of σ^2 in the FE model.

So far, we have assumed that X does not contain a vector of ones. When, in addition to X , there is an intercept δ in the regression, we have

$$e = My = M\iota_N\delta + M\varepsilon = \iota_N\delta + \varepsilon \quad (4.13)$$

[cf. the line below (4.7)], so in that case we will use the centered residuals $f \equiv E_N e$ rather than e . Computing the expectation of the redefined q_H and q_T (with f substituted for e) is essentially the same but is more complicated. See appendix A for results.

When the effects α_h and γ_i are normally distributed, explicit expressions for the variance of the QUE's of σ_1^2 and σ_2^2 can be derived. This is indicated in appendix B.

5. ML estimation and the information matrix

For complete data, ML estimation in error-components models has been studied by Amemiya (1971). Applying the general results obtained by Magnus (1978), the first-order conditions for ML estimation of β and the parameters in Ω are

$$\hat{\beta} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y, \quad (5.1)$$

$$\text{tr}(\hat{\Omega}_\theta^{-1}\Omega) = e'\hat{\Omega}_\theta^{-1}e, \quad \theta = \sigma^2, \sigma_1^2, \text{ or } \sigma_2^2, \quad (5.2)$$

with $e \equiv y - X\hat{\beta}$ and $\hat{\Omega}_\theta^{-1}$ being the derivative of Ω^{-1} with respect to θ . In appendix C, it is shown that

$$\Omega_{\sigma^2}^{-1} = \sigma^{-2} \{ -\Omega^{-1} + \sigma_1^{-2} R \tilde{A}_H^{-2} R' + \sigma_2^{-2} V Z_2 \tilde{Q}^{-2} Z_2' V \}, \quad (5.3)$$

$$\Omega_{\sigma_1^2}^{-1} = -\sigma_1^{-4} R \tilde{A}_H^{-2} R', \quad (5.4)$$

$$\Omega_{\sigma_2^2}^{-1} = -\sigma_2^{-4} V Z_2 \tilde{Q}^{-2} Z_2' V, \quad (5.5)$$

with

$$R \equiv (I_N - V Z_2 \tilde{Q}^{-1} Z_2') Z_1, \quad (5.6)$$

and that

$$\text{tr}(\Omega_{\sigma^2}^{-1}\Omega) = \sigma^{-2} \{ -N + p_1 + p_2 \}, \quad (5.7)$$

$$\text{tr}(\Omega_{\sigma_i^2}^{-1}\Omega) = -\sigma_i^{-2} p_i, \quad i = 1, 2, \quad (5.8)$$

where

$$p_1 \equiv H - \frac{\sigma^2}{\sigma_1^2} \text{tr}(\tilde{A}_H^{-1} + \tilde{A}_H^{-1} A' \tilde{Q}^{-1} A \tilde{A}_H^{-1}), \quad (5.9)$$

$$p_2 \equiv T - \frac{\sigma^2}{\sigma_2^2} \text{tr} \tilde{Q}^{-1}. \quad (5.10)$$

A few comments on these results are in order. When complete data are available, there do not exist closed-form expressions for the MLE's of the variance components, and this holds true a fortiori for the incomplete-data case.

Again, the formulae do not look particularly attractive. Yet it is simple to see that all expressions of interest can be computed in $O(H)$ time. [Since the number of iterations is unknown, this does of course not guarantee that the computation of the MLE's until convergence also can be done in $O(H)$ time.]

Both for the purpose of computing asymptotic standard errors of the ML estimators and to test hypotheses it is useful to compute the information matrix. The information matrix is derived in appendix C to be

$$\begin{bmatrix} X'\Omega^{-1}X & 0 \\ 0 & \frac{1}{2}\psi \end{bmatrix} [(k+3) \times (k+3)], \quad (5.11)$$

with

$$\psi \equiv \begin{bmatrix} \psi_{00} & \psi_{01} & \psi_{02} \\ \psi_{10} & \psi_{11} & \psi_{12} \\ \psi_{20} & \psi_{21} & \psi_{22} \end{bmatrix}, \quad (5.12)$$

and

$$\psi_{00} \equiv \sigma^{-4} \{ N + q_{11} + q_{22} - 2(p_1 + p_2 - q_{12}) \}, \quad (5.13)$$

$$\psi_{0i} = \psi_{i0} \equiv \sigma^{-2} \sigma_i^{-2} (p_i - q_{ii} - q_{12}), \quad i = 1, 2, \quad (5.14)$$

$$\psi_{ij} = \sigma_i^{-2} \sigma_j^{-2} q_{ij}, \quad i, j = 1, 2, \quad (5.15)$$

$$q_{11} \equiv \text{tr} \left(I_H - \frac{\sigma^2}{\sigma_1^2} (\tilde{\Delta}_H^{-1} + \tilde{\Delta}_H^{-1} A' \tilde{Q}^{-1} A \tilde{\Delta}_H^{-1}) \right)^2, \quad (5.16)$$

$$q_{22} \equiv \text{tr} \left(I_T - \frac{\sigma^2}{\sigma_2^2} \tilde{Q}^{-1} \right)^2, \quad (5.17)$$

$$q_{21} = q_{12} \equiv \frac{\sigma^4}{\sigma_1^2 \sigma_2^2} \text{tr} (\tilde{\Delta}_H^{-1} A' \tilde{Q}^{-2} A \tilde{\Delta}_H^{-1}). \quad (5.18)$$

All these expressions can be computed in $O(H)$ time.

The problem of estimation with incomplete panels can be considered as a 'missing-observations' problem. By 'missing' we mean that for some (h, t) pairs both y_{ht} and x_{ht} are not observed. In the context of missing observa-

tions, the EM algorithm for obtaining ML estimates has attracted a lot of attention, recently also in econometrics. Fair's treatment of the Tobit model is possibly the best-known example [Fair (1977)]. In our context the EM approach suggests the following. Choose starting values for the model parameters, and compute the distribution of the 'missing' y_{ht} (with $x_{ht} = 0$ without loss of generality) conditioned on the observed data and the starting values. Next write down the likelihood function for the complete panel (with its nicely structured Ω), take its expectation with respect to the missing ' y_{ht} ', and maximize it to obtain new parameter values, etc. As far as we know, the EM algorithm has not yet been elaborated for the ECM with incomplete observations. It could offer a useful alternative to our approach, but a preliminary inspection suggests that working out the EM approach produces a lot of messy algebra, much of the same type as with our approach. This apparently is inherent to the problem at hand.

6. Some simulation results

We consider the following simple version of model (2.1): $\sigma_1^2 = \text{var}(\alpha_h) = 400$, $\sigma_2^2 = \text{var}(\gamma_t) = 25$, $\sigma^2 = \text{var}(u_{ht}) = 25$, $x_{ht} = (1, \tilde{x}'_{ht})'$, $\beta = (\beta_0, \beta_1)' = (25, 2)'$, $H = 100$, $T = 5$. The scalars x_{ht} were generated according to the scheme introduced by Nerlove (1971a, p. 367) and subsequently used by, e.g., Arora (1973), Baltagi (1981), and Heckman (1981):

$$x_{ht} = 0.1t + 0.5x_{h,t-1} + \omega_{ht}, \quad (6.1)$$

with the ω_{ht} uniform $[-\frac{1}{2}, \frac{1}{2}]$ and $x_{h0} = 5 + 10\omega_{h0}$. Three cases are considered:

1. No observations are missing: 'complete data'.
2. Each period 20% of the households left in the panel is removed randomly: 'random attrition'.
3. In period 1 we start with 40 households. In period 2, 20 new households are added. In period 3, 20 households remaining from period 1 are removed and 20 new households are added. In period 4, the 20 households still remaining from period 1 are removed and 20 new households are added, etc.: 'rotating panel'.

The x values have been drawn once and are used in all experiments. Given the x values we generated values of y_{ht} according to model (2.1) in each new simulation run. For each case 50 runs are made, all using the same pattern of

Table 1
Computational burden of different estimation methods.^a

Set up	OLS	GLS	ML
Complete data	100	414	2026
Random attrition	68	333	3296
Rotating panel	53	297	3663

^aIn CPU seconds relative to OLS on the complete data. The computing time for ML includes the time required to generate starting values by means of the GLS procedure.

'missings' in the 'random attrition' case. For the ML estimation we always use the estimates of the two-step GLS procedure as starting values.

Table 1 gives an indication of the average computing time required by the two estimation methods for each case (GLS and ML), plus OLS. It is quite obvious that GLS is a lot cheaper than ML. Furthermore, computation time for GLS *decreases* if more observations are missing, whereas computation time for ML *increases*. For complete data (100 households each period) ML requires about 5 times more CPU seconds than GLS. In the rotating panel case (40 or 60 households each period), ML takes about 12 times more CPU seconds than GLS. Of course, the precise magnitude of these figures depend on the particular design matrix chosen.

Table 2 presents means and variances of various estimates obtained in the 50 simulation runs. After each figure, its variance over the 50 runs is given. As to the regression coefficients, it appears that OLS, though unbiased conditional on the x_h , performs badly, and that GLS and ML give nearly identical results, both in terms of point estimates and of sampling variances. The latter are somewhat smaller than those obtained with FE. The sampling variances of β_1 increase over the three cases, which are ordered by a decreasing number of observations, and go markedly up when moving to the 'rotating panel' case with its short time series.

Regarding the estimation of the variance components, it is striking that on average the unbiased QUE's are in all cases at least as close to the true value as the MLE's, and in a majority of cases much better. If one is interested in the values of the variance components, the results suggest a preference for the QUE's – iteration does not seem to pay off.

The column 'variance of GLS estimate' gives the means of the estimates of the variance of the estimates of σ_1^2 , σ_2^2 , and σ^2 , respectively. Although the variance formulae indicated in appendix B are exact, the estimates actually used are not unbiased, because we have to plug in estimates of σ^2 , σ_1^2 , and σ_2^2 . The 'variance of the ML estimate' is based on the information matrix. It turns out that the variance estimates for GLS and ML show some agreement, both between the two and with the corresponding sampling variances, but there are some striking exceptions – variances of variances can be volatile.

Table 2
Simulation results.^a

Parameter	True value	OLS estimate	FE estimate	GLS estimate	Variance of GLS estimate	ML estimate	Variance of ML estimate
<i>Complete data</i>							
β_0	25	27.5 (139)	0 (0)	26.0 (90.1)		26.0 (90.1)	
β_1	2	0.973 (24.2)	1.97 (0.122)	1.95 (0.109)		1.95 (0.109)	
σ_1^2	400			389 (109000)	$130000 (1.37 \times 10^{11})$	311 (69800)	$66800 (3.61 \times 10^{10})$
σ_2^2	25			25.3 (15.1)	3.00 (199)	25.3 (14.9)	18.9 (22.7)
σ^2	25			24.8 (2.57)	3.12 (0.162)	24.8 (2.57)	3.12 (0.162)
<i>Random attrition (20%)</i>							
β_0	25	26.9 (135)	0 (0)	26.0 (91.8)		26.0 (92.1)	
β_1	2	1.12 (22.2)	1.96 (0.264)	1.93 (0.160)		1.94 (0.168)	
σ_1^2	400			386 (119000)	$144000 (2.25 \times 10^{11})$	311 (70100)	$66800 (3.66 \times 10^{10})$
σ_2^2	25			23.5 (290)	$643 (4.27 \times 10^6)$	21.6 (94.7)	21.7 (88.2)
σ^2	25			24.9 (4.63)	5.34 (0.869)	28.3 (75.5)	7.26 (25.7)
<i>Rotating panel</i>							
β_0	25	27.6 (167)	0 (0)	26.1 (89.1)		26.2 (89.6)	
β_1	2	0.851 (27.6)	1.94 (0.587)	1.88 (0.415)		1.88 (0.407)	
σ_1^2	400			390 (108000)	$134000 (1.44 \times 10^{11})$	312 (70100)	$67100 (3.52 \times 10^{10})$
σ_2^2	25			22.3 (419)	$5750 (2.55 \times 10^8)$	17.7 (136)	21.7 (64.4)
σ^2	25			24.9 (7.74)	7.97 (3.16)	31.7 (139)	14.5 (123)

^a Entries are means over 50 simulation runs; numbers in parentheses are corresponding variances.

7. Concluding remarks

Altogether, missing observations lead to less elegant expressions for estimators, and computer programs are accordingly more complicated. Yet, in terms of computational complexity missing observations do not constitute major problems. In view of the large difference in computational cost, GLS is to be preferred over ML in the ECM. Statistically, ML and GLS do not seem to behave very differently in finite samples. In the ECM, the FE estimator is also a viable alternative. It has a somewhat larger sampling variance, but one has not to make the assumption that the random effects are independent from the regressors, which may be troublesome in many applications.

Appendix A: QUE's in a model with an intercept

When an intercept term is present, we develop the QUE's starting from quadratic functions of the centered residuals:

$$q_H' \equiv f' Z_2 \Delta_T^{-1} Z_2' f, \quad (\text{A.1})$$

$$q_T' \equiv f' Z_1 \Delta_H^{-1} Z_1' f. \quad (\text{A.2})$$

In this appendix we evaluate the expectation of q_H' , which makes clear how unbiased estimators of σ_1^2 and σ_2^2 can be constructed. As $f \equiv E_N e = E_N M \epsilon$,

$$q_H' = \epsilon' M' E_N Z_2 \Delta_T^{-1} Z_2' E_N M \epsilon \quad (\text{A.3})$$

and, by elaborating E_N ,

$$\begin{aligned} E(q_H') &= E(q_H) + \iota_N' Z_2 \Delta_T^{-1} Z_2' \iota_N \iota_N' M \Omega M' \iota_N / N^2 \\ &\quad - 2 \iota_N' M \Omega M' Z_2 \Delta_T^{-1} Z_2' \iota_N / N. \end{aligned} \quad (\text{A.4})$$

As $Z_2 \Delta_T^{-1} Z_2' \iota_N = \iota_N$, this carries over into

$$\begin{aligned} E(q_H') &= E(q_H) - \iota_N' M \Omega M' \iota_N / N \\ &= E(q_H) - \{ \iota_N' \Omega \iota_N + \sigma^2 \iota_N' X (X' P X)^{-1} X' \iota_N \} / N \\ &= E(q_H) - \sigma^2 (1 + k_0) - (\sigma_1^2 \lambda_1 + \sigma_2^2 \lambda_2) / N, \end{aligned} \quad (\text{A.5})$$

where the second equality is based on (4.9) and $Z_1 \Delta_T^{-1} Z_1' \iota_N = \iota_N$ and where in

the last step the following definitions have been used:

$$k_0 \equiv \iota'_N X (X' P X)^{-1} X' \iota_N / N, \quad (\text{A.6})$$

$$\lambda_1 \equiv \iota'_N Z_1 Z_1' \iota_N \equiv \sum_{h=1}^H m_h^2, \quad (\text{A.7})$$

$$\lambda_2 \equiv \iota'_N Z_2 Z_2' \iota_N \equiv \sum_{t=1}^T n_t^2. \quad (\text{A.8})$$

The interpretation of m_h is that it is the number of times ($2 \leq m_h \leq T$) that household h was in the panel, whereas n_t ($2 \leq n_t \leq H$) denotes the number of households in the panel at time t .

So the presence of an intercept term introduces an adjustment of $E(q_H)$; see the last line of (A.5). The same adjustment has to be made to $E(q_T)$.

Appendix B: On the variances of the QUE's

Let again θ denote any one of the variance components σ^2 , σ_1^2 , or σ_2^2 . Their QUE's can be written as

$$\hat{\theta} = \varepsilon' M E_N W E_N M' \varepsilon = f' W f, \quad (\text{B.1})$$

with

$$W = aP + bZ_1 \Delta_H^{-1} Z_1' + cZ_2 \Delta_T^{-1} Z_2'. \quad (\text{B.2})$$

Here a , b , and c are constants that can be chosen such that (B.1)–(B.2) generate the QUE's. According to multivariate normal theory,

$$\text{var}(\hat{\theta}) = 2 \text{tr}(M E_N W E_N M' \Omega)^2. \quad (\text{B.3})$$

Elaborating (B.3) is a tedious affair but is in principle straightforward. We have elaborated and programmed (B.3) to compute the variances of the QUE's in the simulations reported in section 6. Since the formulae for the variances are ugly and do not yield any insights, they are not given here. The formulae are available from the authors on request.

Appendix C: Derivation of results on ML

Let θ denote any one of the parameters σ^2 , σ_1^2 , or σ_2^2 . Then

$$\begin{aligned} \Omega_\theta^{-1} &= \sigma^{-2} \{ -\sigma_\theta^2 \Omega^{-1} + (V - V Z_2 \tilde{Q}^{-1} Z_2' V)_\theta \} \\ &= \sigma^{-2} \{ -\sigma_\theta^2 \Omega^{-1} + V_\theta - V_\theta Z_2 \tilde{Q}^{-1} Z_2' V \\ &\quad + V Z_2 \tilde{Q}^{-1} \tilde{Q} \tilde{Q}^{-1} Z_2' V - V Z_2 \tilde{Q}^{-1} Z_2' V_\theta \}. \end{aligned} \quad (\text{C.1})$$

There holds

$$\tilde{Q}_\theta = \tilde{\Delta}_{T\theta} + A\tilde{\Delta}_H^{-1}\Delta_{H\theta}\tilde{\Delta}_H^{-1}A' = \tilde{\Delta}_{T\theta} + Z_2'V_\theta Z_2, \quad (C.2)$$

so

$$VZ_2\tilde{Q}^{-1}\tilde{Q}_\theta\tilde{Q}^{-1}Z_2'V = VZ_2\tilde{Q}^{-1}\tilde{\Delta}_{T\theta}\tilde{Q}^{-1}Z_2'V + VZ_2\tilde{Q}^{-1}Z_2'V_\theta Z_2\tilde{Q}^{-1}Z_2'V. \quad (C.3)$$

Substitution of (C.3) into (C.1) yields

$$\begin{aligned} \Omega_\theta^{-1} = \sigma^{-2} \{ & -\sigma_\theta^2 \Omega^{-1} + (I_N - VZ_2\tilde{Q}^{-1}Z_2')V_\theta(I_N - Z_2\tilde{Q}^{-1}Z_2'V) \\ & + VZ_2\tilde{Q}^{-1}\tilde{\Delta}_{T\theta}\tilde{Q}^{-1}Z_2'V \}. \end{aligned} \quad (C.4)$$

In view of the definition of V [see (3.3)], there holds

$$V_\theta = Z_1\tilde{\Delta}_H^{-1}\tilde{\Delta}_{H\theta}\tilde{\Delta}_H^{-1}Z_1'. \quad (C.5)$$

This expression and the definition of R [see (5.6)] allow for writing (C.4) as

$$\Omega_\theta^{-1} = \sigma^{-2} \{ -\sigma_\theta^2 \Omega^{-1} + R\tilde{\Delta}_H^{-1}\tilde{\Delta}_{H\theta}\tilde{\Delta}_H^{-1}R' + VZ_2\tilde{Q}^{-1}\tilde{\Delta}_T\tilde{Q}^{-1}Z_2'V \}. \quad (C.6)$$

It remains to substitute for σ_θ^2 , $\tilde{\Delta}_{T\theta}$, and $\tilde{\Delta}_{H\theta}$ in order to establish (5.3)–(5.5). There holds

$$\tilde{\Delta}_{T\sigma^2} = \sigma_2^{-2}I_T, \quad \tilde{\Delta}_{T\sigma_1^2} = 0, \quad \tilde{\Delta}_{T\sigma_2^2} = -\frac{\sigma^2}{\sigma_2^4}I_T, \quad (C.7)$$

$$\tilde{\Delta}_{H\sigma^2} = \sigma_1^{-2}I_H, \quad \tilde{\Delta}_{H\sigma_1^2} = -\frac{\sigma^2}{\sigma_1^4}I_H, \quad \tilde{\Delta}_{H\sigma_2^2} = 0, \quad (C.8)$$

and this leads to (5.3)–(5.5) directly.

For further results, we need an expression for $\Omega_\theta^{-1}\Omega$. In view of (C.6), this means that we first have to consider $R'\Omega$ and $Z_2'V\Omega$. Now, since

$V^{-1} = I_N + (\sigma_1/\sigma)^2 Z_1 Z_1'$, $\Omega = \sigma^2 V^{-1} + \sigma_2^2 Z_2 Z_2'$. Hence

$$\begin{aligned}
 Z_2' V \Omega &= Z_2' V (\sigma^2 V^{-1} + \sigma_2^2 Z_2 Z_2') \\
 &= \sigma^2 Z_2' + \sigma_2^2 Z_2' V Z_2 Z_2' \\
 &= \sigma^2 Z_2' + \sigma_2^2 \left(\tilde{Q} - \frac{\sigma^2}{\sigma_2^2} I_T \right) Z_2' \\
 &= \sigma^2 Z_2' + \sigma_2^2 \tilde{Q} Z_2' - \sigma^2 Z_2' \\
 &= \sigma_2^2 \tilde{Q} Z_2'
 \end{aligned} \tag{C.9}$$

and

$$\begin{aligned}
 R' \Omega &= Z_1' (I_N - Z_2 Q^{-1} Z_2' V) \Omega \\
 &= Z_1' \Omega - Z_1' Z_2 \tilde{Q}^{-1} Z_2' V \Omega \\
 &= Z_1' \Omega - \sigma_2^2 Z_1' Z_2 \tilde{Q}^{-1} \tilde{Q} Z_2' \\
 &= \sigma^2 Z_1' + \sigma_2^2 Z_1' Z_2 Z_2' + \sigma_1^2 Z_1' Z_1 Z_1' - \sigma_2^2 Z_1' Z_2 Z_2' \\
 &= \sigma^2 Z_1' + \sigma_1^2 \left(\tilde{\Delta}_H - \frac{\sigma^2}{\sigma_1^2} I_H \right) Z_1' \\
 &= \sigma_1^2 \tilde{\Delta}_H Z_1'.
 \end{aligned} \tag{C.10}$$

These expressions, with (5.3)–(5.5), yield

$$(\partial \Omega^{-1} / \partial \sigma^2) \Omega = \sigma^{-2} \{ -I_N + R \tilde{\Delta}_H^{-1} Z_1' + V Z_2 \tilde{Q}^{-1} Z_2' \}, \tag{C.11}$$

$$(\partial \Omega^{-1} / \partial \sigma_1^2) \Omega = -\sigma_1^{-2} R \tilde{\Delta}_H^{-1} Z_1', \tag{C.12}$$

$$(\partial \Omega^{-1} / \partial \sigma_2^2) \Omega = -\sigma_2^{-2} V Z_2 \tilde{Q}^{-1} Z_2'. \tag{C.13}$$

To verify (5.7) and (5.8) we have to take traces of these expressions. We use

the following facts:

$$\begin{aligned}
 \text{tr}(R\tilde{\Delta}_H^{-1}Z_1') &= \text{tr}(Z_1'R\tilde{\Delta}_H^{-1}) \\
 &= \text{tr}Z_1'(I_N - VZ_2\tilde{Q}^{-1}Z_2')Z_1\tilde{\Delta}_H^{-1} \\
 &= \text{tr}(Z_1'Z_1 - Z_1'VZ_2\tilde{Q}^{-1}A)\tilde{\Delta}_H^{-1} \\
 &= \text{tr}\left(Z_1'Z_1 - \frac{\sigma^2}{\sigma_1^2}\tilde{\Delta}_H^{-1}A'\tilde{Q}^{-1}A\right)\tilde{\Delta}_H^{-1} \\
 &= \text{tr}\left(\tilde{\Delta}_H - \frac{\sigma^2}{\sigma_1^2}I_H - \frac{\sigma^2}{\sigma_1^2}\tilde{\Delta}_H^{-1}A'\tilde{Q}^{-1}A\right)\tilde{\Delta}_H^{-1} \\
 &= H - \frac{\sigma^2}{\sigma_1^2}\text{tr}(\tilde{\Delta}_H^{-1} + \tilde{\Delta}_H^{-1}A'\tilde{Q}^{-1}A\tilde{\Delta}_H^{-1}) \\
 &= p_1 \quad [\text{cf. (5.9)}].
 \end{aligned} \tag{C.14}$$

$$\begin{aligned}
 \text{tr}(VZ_2\tilde{Q}^{-1}Z_2') &= \text{tr}(Z_2'VZ_2\tilde{Q}^{-1}) \\
 &= \text{tr}\left(\tilde{Q} - \frac{\sigma^2}{\sigma_2^2}I_T\right)\tilde{Q}^{-1} \\
 &= T - \frac{\sigma^2}{\sigma_2^2}\text{tr}\tilde{Q}^{-1} \\
 &= p_2 \quad [\text{cf. (5.10)}].
 \end{aligned} \tag{C.15}$$

This establishes (5.7) and (5.8).

We finally derive the elements of the information matrix (5.12). The element of ψ [cf. (5.12)] corresponding with parameters θ and θ' , say, is $\text{tr}(\Omega_\theta^{-1}\Omega\Omega_{\theta'}^{-1}\Omega)$, and is hence obtained by taking the trace of the product of any two right-hand

sides of (C.11)–(C.13). The following facts are used

$$\begin{aligned}\text{tr}(R\tilde{\Delta}_H^{-1}Z_1'R\tilde{\Delta}_H^{-1}Z_1') &= \text{tr}\left(I_H - \frac{\sigma^2}{\sigma_1^2}(\tilde{\Delta}_H^{-1} + \tilde{\Delta}_H^{-1}A'\tilde{Q}^{-1}A\tilde{\Delta}_H^{-1})\right)^2 \\ &= q_{11} \quad [\text{cf. (5.16)}],\end{aligned}\tag{C.16}$$

$$\begin{aligned}\text{tr}(VZ_2\tilde{Q}^{-1}Z_2'VZ_2\tilde{Q}^{-1}Z_2') &= \text{tr}(Z_2'VZ_2\tilde{Q}^{-1})^2 \\ &= \text{tr}\left(I_T - \frac{\sigma^2}{\sigma_2^2}\tilde{Q}^{-1}\right)^2 \\ &= q_{22} \quad [\text{cf. (5.17)}],\end{aligned}\tag{C.17}$$

$$\begin{aligned}\text{tr}(VZ_2\tilde{Q}^{-1}Z_2'R\tilde{\Delta}_H^{-1}Z_1') &= \text{tr}(Z_1'VZ_2\tilde{Q}^{-1}Z_2'R\tilde{\Delta}_H^{-1}) \\ &= \frac{\sigma^4}{\sigma_1^2\sigma_2^2} \text{tr}(\tilde{\Delta}_H^{-1}A'\tilde{Q}^{-2}A\tilde{\Delta}_H^{-1}) \\ &= q_{12} \quad [\text{cf. (5.18)}],\end{aligned}\tag{C.18}$$

where the second equality sign in (C.18) is based on

$$\begin{aligned}Z_2'R &= Z_2'(I_N - VZ_2\tilde{Q}^{-1}Z_2')Z_1 \\ &= A - Z_2'VZ_2\tilde{Q}^{-1}A \\ &= \left(I_T - \left(\tilde{Q} - \frac{\sigma^2}{\sigma_2^2}I_T\right)\tilde{Q}^{-1}\right)A \\ &= \frac{\sigma^2}{\sigma_2^2}\tilde{Q}^{-1}A.\end{aligned}\tag{C.19}$$

With the aid of (C.16)–(C.18), the elements of the information matrix follow directly.

References

- Amemiya, T., 1971, The estimation of the variances in a variance-components model, *International Economic Review* 12, 1-13.
- Arora, S., 1973, Error components regression models and their applications, *Annals of Economic and Social Measurement* 2, 451-462.
- Balestra, P., 1973, Best quadratic unbiased estimators of the variance-covariance matrix in normal regression, *Journal of Econometrics* 1, 17-28.
- Balestra, P. and M. Nerlove, 1966, Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas, *Econometrica* 34, 585-612.
- Baltagi, B.H., 1981, Pooling: An experimental study of alternative testing and estimation procedures in a two-way error component model, *Journal of Econometrics* 17, 21-49.
- Baltagi, B.H., 1985, Pooling cross-sections with unequal time-series lengths, *Economics Letters* 18, 133-136.
- Biørn, E., 1981, Estimating economic relations from incomplete cross-section/time-series data, *Journal of Econometrics* 16, 221-236.
- Dielman, T.E., 1983, Pooled cross-sectional and time series data: A survey of current statistical methodology, *The American Statistician* 37, 111-122.
- Fair, R.C., 1977, A note on the computation of the Tobit estimator, *Econometrica* 45, 1723-1728.
- Fuller, W.A. and G.E. Battese, 1974, Estimation of linear functions with crossed-error structure, *Journal of Econometrics* 2, 67-78.
- Hausman, J.A. and D.A. Wise, 1977, Social experimentation, truncated distributions, and efficient estimation, *Econometrica* 45, 919-938.
- Heckman, J.J., 1981, The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process, in: C. Manski and D. McFadden, eds., *Structural analysis of discrete data with econometric applications* (MIT Press, Cambridge, MA).
- Hsiao, C., 1986, *Analysis of panel data* (Cambridge University Press, Cambridge).
- Magnus, J.R., 1978, Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix, *Journal of Econometrics* 7, 281-312.
- Mazodier, P., 1972, L'estimation des modèles à erreurs composées, *Annales de l'INSEE* 7, 43-72.
- Mazodier, P., ed., 1978, *The econometrics of panel data*, Special Edition of *Annales de l'INSEE*, 30-31.
- Nerlove, M., 1971a, Further evidence on the estimation of dynamic economic relations from a time series of cross sections, *Econometrica* 39, 359-382.
- Nerlove, M., 1971b, A note on error components models, *Econometrica* 39, 383-396.
- Taylor, W.E., 1977, Small sample properties of a class of two stage Aitken estimators, *Econometrica* 45, 497-508.
- Wallace, T.D. and A. Hussain, 1969, The use of error component models in combining cross section with time series data, *Econometrica* 37, 55-72.
- Wansbeek, T.J. and A. Kapteyn, 1982, A class of decompositions of the variance covariance matrix of a generalized error components model, *Econometrica* 50, 713-724.

Reprint Series, CentER, Tilburg University, The Netherlands:

- No. 1 G. Marini and F. van der Ploeg, Monetary and fiscal policy in an optimising model with capital accumulation and finite lives, The Economic Journal, Vol. 98, No. 392, 1988, pp. 772 - 786.
- No. 2 F. van der Ploeg, International policy coordination in interdependent monetary economies, Journal of International Economics, Vol 25, 1988, pp. 1 - 23.
- No. 3 A.P. Barten, The history of Dutch macroeconomic modelling (1936-1986), in W. Driehuis, M.M.G. Fase and H. den Hartog (eds.), Challenges for Macroeconomic Modelling, Contributions to Economic Analysis 178, Amsterdam: North-Holland, 1988, pp. 39 - 88.
- No. 4 F. van der Ploeg, Disposable income, unemployment, inflation and state spending in a dynamic political-economic model, Public Choice, Vol. 60, 1989, pp. 211 - 239.
- No. 5 Th. ten Raa and F. van der Ploeg, A statistical approach to the problem of negatives in input-output analysis, Economic Modelling, Vol. 6, No. 1, 1989, pp. 2 - 19.
- No. 6 E. van Damme, Renegotiation-proof equilibria in repeated prisoners' dilemma, Journal of Economic Theory, Vol. 47, No. 1, 1989, pp. 206 - 217.
- No. 7 C. Mulder and F. van der Ploeg, Trade unions, investment and employment in a small open economy: a Dutch perspective, in J. Muysken and C. de Neubourg (eds.), Unemployment in Europe, London: The MacMillan Press Ltd, 1989, pp. 200 - 229.
- No. 8 Th. van de Klundert and F. van der Ploeg, Wage rigidity and capital mobility in an optimizing model of a small open economy, De Economist 137, nr. 1, 1989, pp. 47 - 75.
- No. 9 G. Dhaene and A.P. Barten, When it all began: the 1936 Tinbergen model revisited, Economic Modelling, Vol. 6, No. 2, 1989, pp. 203 - 219.
- No. 10 F. van der Ploeg and A.J. de Zeeuw, Conflict over arms accumulation in market and command economies, in F. van der Ploeg and A.J. de Zeeuw (eds.), Dynamic Policy Games in Economics, Amsterdam: Elsevier Science Publishers B.V. (North-Holland), 1989, pp. 91 - 119.
- No. 11 J. Driffill, Macroeconomic policy games with incomplete information: some extensions, in F. van der Ploeg and A.J. de Zeeuw (eds.), Dynamic Policy Games in Economics, Amsterdam: Elsevier Science Publishers B.V. (North-Holland), 1989, pp. 289 - 322.
- No. 12 F. van der Ploeg, Towards monetary integration in Europe, in P. De Grauwe e.a., De Europese Monetaire Integratie: vier visies, Wetenschappelijke Raad voor het Regeringsbeleid V 66, 's-Gravenhage: SDU uitgeverij, 1989, pp. 81 - 106.

- No. 13 R.J.M. Alessie and A. Kapteyn, Consumption, savings and demography, in A. Wenig, K.F. Zimmermann (eds.), Demographic Change and Economic Development, Berlin/Heidelberg: Springer-Verlag, 1989, pp. 272 - 305.
- No. 14 A. Hoque, J.R. Magnus and B. Pesaran, The exact multi-period mean-square forecast error for the first-order autoregressive model, Journal of Econometrics, Vol. 39, No. 3, 1988, pp. 327 - 346.
- No. 15 R. Alessie, A. Kapteyn and B. Melenberg, The effects of liquidity constraints on consumption: estimation from household panel data, European Economic Review 33, No. 2/3, 1989, pp. 547 - 555.
- No. 16 A. Holly, J.R. Magnus, A note on instrumental variables and maximum likelihood estimation procedures, Annales d'Économie et de Statistique, No. 10, April-June, 1988, pp. 121 - 138.
- No. 17 P. ten Hacken, A. Kapteyn and I. Woittiez, Unemployment benefits and the labor market, a micro/macro approach, in B.A. Gustafsson and N. Anders Klevmarken (eds.), The Political Economy of Social Security, Contributions to Economic Analysis 179, Amsterdam: Elsevier Science Publishers B.V. (North-Holland), 1989, pp. 143 - 164.
- No. 18 T. Wansbeek and A. Kapteyn, Estimation of the error-components model with incomplete panels, Journal of Econometrics Vol. 41, No. 3, 1989, pp. 341 - 361.

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS

Bibliotheek K. U. Brabant

17 000 01169683 9