A class of decompositions of the variance covariance matrix of a generalized error components model

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

Published in:
Econometrica

Publication date:
1982

Econometrica, Vol. 50, No. 3 (May, 1982)

A CLASS OF DECOMPOSITIONS OF THE VARIANCE-COVARIANCE MATRIX OF A GENERALIZED ERROR COMPONENTS MODEL[1]

BY TOM WANSBEEK AND ARIE KAPTEYN

A class of decompositions is derived for the variance-covariance matrix Ω of a generalized error components model, introduced in [18 and 19]. The spectral decomposition of Ω is a member of this class. For estimation purposes certain other members of the class are preferred, especially those that allow for simplifying transformations of the model not depending on unknown parameters. The transformations suggest simple and asymptotically efficient estimators of both the parameters in Ω and the parameters in the systematic part of the model.

1. INTRODUCTION

THE ERROR COMPONENTS (EC) model has become a popular means of analysis of panel data, i.e., a time series of cross sections. Since the first papers by, among others, Hildreth [3], Kuh [6], and Balestra and Nerlove [2], a steady stream of theoretical and empirical papers has emerged, culminating in a special conference on the econometrics of panel data in 1977 (Mazodier [9]).

The EC model can be written as follows:

$$y_{nt} = \sum_{j=1}^{k} \beta_j x_{ntj} + \mu_n + \nu_t + u_{nt}, \tag{1}$$

where n refers to the unit of observation in the cross section, $n = 1, \ldots, N$, and t is the time index of the observation, $t = 1, \ldots, T$. The β_j are fixed parameters and the u_{nt} are i.i.d. error terms. The μ_n and ν_t are i.i.d. random variables, mutually independent and independent of x_{ntj} and u_{nt}. An alternative assumption would be that μ_n and ν_t are fixed unknown constants, which transforms (1) into an analysis of covariance model. The connection between the analysis of covariance model and the EC model has been studied extensively by Mundlak [11].

Various generalizations and variations of (1) are possible. For example, analogous to having the error term consisting of three components μ_n, ν_t, u_{nt}, one may as well decompose β into three parts, i.e., $\beta_{ntj} = \beta_j + \mu_{nj} + \nu_{tj}$. This has been done by Hsiao [4, 5] and Swamy and Mehta [15], among others. If the model contains a constant term, then the error term has again three components. The case where μ_{nj} and ν_{tj} are fixed has been studied by Malinvaud and Milleron [8]. Another generalization has been proposed by Mazodier and Trognon [10], who allow for heteroskedasticity of μ_n and ν_t. They derive estimators of β and the variances of

[1] The paper was written while the first author was with the Department of Economics, University of Southern California, and with the Center for Research in Public Economics, Leyden University. We thank H. Neudecker and an anonymous referee for their detailed comments.

the error components for the case where there are "strata" (i.e., subsets of the domain of n and t) for which the variances of μ_n and ν_t are constant.

In two earlier papers [19, 20] we have considered the following generalization of (1):

$$y_{nt} = \sum_{j=1}^{k} \beta_j x_{ntj} + \sum_{j=1}^{p} \alpha_{1tj} z_{1nj} + \sum_{j=1}^{q} \alpha_{2nj} z_{2tj} + u_{nt}. \tag{2}$$

Evidently, (2) reduces to (1) if $p = q = 1$ and z_{1nj} and z_{2tj} are identically equal to one. Just like μ_n and ν_t in (1) we can take α_{1tj} and α_{2nj} to be either fixed or random. The case of fixed α_{1tj} and α_{2nj} has been analyzed in [19]. Here we take α_{1tj} and α_{2nj} to be random (more specific assumptions are given below).

Model (2) is a special case of the model considered by Hsiao. By concentrating on a more special case we are better able to exploit its mathematical structure. Model (2) is still sufficiently general to be used in a variety of applications. In [19, 20] we have suggested various applications for (2), like the estimation of production functions or household expenditure functions. The model is also applicable outside a time series of cross sections context, e.g., in the analysis of traffic flows. An empirical application is given in [18, Section 3.5].

Both from a practical and a theoretical perspective, model (2) has some properties to commend it. This is best brought out by first reviewing a few results that have emerged in the error components literature.

Let the variance-covariance matrix of the composite error term $\mu_n + \nu_t + u_{nt}$ of model (1) be Φ, defined as

$$\Phi \equiv \sigma^2 I_{NT} + \sigma_\nu^2 I_T \otimes J_N + \sigma_\mu^2 J_T \otimes I_N, \tag{3}$$

where I denotes the unit matrix of order indicated by its subscript, J is a square matrix of unit elements of order indicated by its subscript, σ^2, σ_μ^2, σ_ν^2 are the respective variances of u_{nt}, μ_n, ν_t. The symbol "$\otimes$" denotes a Kronecker product. The obvious way to estimate the parameters β_j is by means of GLS, which requires an estimate of Φ and an expression for its inverse (also ML requires Φ^{-1}). In almost any practical application Φ will be far too large to invert it numerically. Wallace and Hussain [17] present a formula for Φ^{-1}, obtained by "trial, error, and generalization." A more systematic analysis was presented by Nerlove [13] who derived the eigenvalues and eigenvectors of Φ, thus allowing for a spectral decomposition of Φ, i.e., Φ is written as:

$$\Phi = \sum_{i=1}^{m} \lambda_i P_i \tag{4}$$

where the λ_i are the m distinct eigenvalues of Φ, of multiplicity r_i, and the P_i are mutually orthogonal, idempotent matrices of rank r_i. The r_i add up to NT, the order of Φ, and the P_i add up to the unit matrix of order NT.

The advantage of knowing the spectral decomposition of Φ is that its powers are easily evaluated, since, for arbitrary γ, there holds:

$$\Phi^{\gamma}=\sum_{i=1}^{m}\lambda_i^{\gamma}P_i. \tag{5}$$

In particular, Φ^{-1} is found by taking $\gamma=-1$. One may also take $\gamma=-\frac{1}{2}$ and use $\Phi^{-1/2}$ to transform the EC-model to a model with an i.i.d. error term for which OLS is the appropriate estimation technique.

A special feature of the spectral decomposition (4) is that only the λ_i appear to depend on the unknown parameters σ^2, σ_μ^2, σ_ν^2. The P_i are known. This suggests that the unknown parameters in Φ can be estimated by first estimating the λ_i. Balestra [1] has derived best quadratic unbiased estimators of the λ_i. His estimator for each λ_i can easily be seen to be the OLS residual variance estimator in the transformed model, obtained by premultiplying (1) by the $r_i \times$ NT-matrix L_i' satisfying $L_i'L_i=I_{r_i}$, $L_iL_i'=P_i$. Once the λ_i are estimated, one can, in principle, derive estimates of σ^2, σ_μ^2, σ_ν^2 from them. Unfortunately, there are four different eigenvalues, which gives rise to different solutions for σ^2, σ_μ^2, σ_ν^2. As will be discussed briefly in Section 3, the optimal solution is to solve σ^2, σ_μ^2, σ_ν^2 from the equations containing the eigenvalues of highest multiplicity. The resulting estimator is asymptotically as efficient as ML or as Rao's MINQUE estimator [14]. Computationally it is very simple compared to the other two. Moreover, the spectral decomposition allows for a ready interpretation of GLS, as given by Nerlove [13].

Mazodier and Trognon [10] have derived somewhat similar results for their generalization. The eigenvalues of highest multiplicity have corresponding eigenvectors that are known, but there is a "remainder" of eigenvalues and eigenvectors that both depend on unknown parameters.

In this paper we derive an analogous result for model (2). It will turn out, however, that there are other decompositions than the spectral decomposition, which have more desirable properties from an estimation viewpoint. By constructing a certain orthonormal basis for the space spanned by the columns of the variance-covariance matrix of the errors, we obtain a class of decompositions of which the spectral decomposition is a member. Another member of the class decomposes the variance-covariance matrix into a known and an unknown part. The decomposition suggests convenient transformations of the model and some simple estimators. The decomposition is closely related to the spectral decomposition, but it has some definite advantages over the latter. Our analysis provides insight into the generality of the EC results reviewed above. For example, it turns out that for the EC model the class of decompositions reduces to one element: the spectral decomposition.

Throughout, we deal with (2) in matrix notation. Define:

$$y\equiv(y_{11},\ldots,y_{N1},\ldots,y_{1t},\ldots,y_{Nt},\ldots,y_{1T},\ldots,y_{NT})', \qquad NT\times 1, \tag{6}$$

$$u\equiv(u_{11},\ldots,u_{N1},\ldots,u_{1t},\ldots,u_{Nt},\ldots,u_{1T},\ldots,u_{NT})', \qquad NT\times 1, \tag{7}$$

(8) $\beta \equiv (\beta_1, \ldots, \beta_k)'$, $k \times 1$,

(9) $\alpha_1 \equiv (\alpha_{111}, \ldots, \alpha_{11p}, \ldots, \alpha_{1T1}, \ldots, \alpha_{1Tp})'$, $Tp \times 1$,

(10) $\alpha_2 \equiv (\alpha_{211}, \ldots, \alpha_{2N1}, \ldots, \alpha_{21q}, \ldots, \alpha_{2Nq})'$, $Nq \times 1$,

(11)
$$X \equiv \begin{bmatrix} x_{111} & \cdots & x_{11k} \\ \vdots & & \vdots \\ x_{N11} & \cdots & x_{N1k} \\ \vdots & & \vdots \\ x_{1T1} & \cdots & x_{1Tk} \\ \vdots & & \vdots \\ x_{NT1} & \cdots & x_{NTk} \end{bmatrix}, \qquad NT \times k,$$

(12)
$$Z_1 \equiv \begin{bmatrix} z_{111} & \cdots & z_{11p} \\ \vdots & & \vdots \\ z_{1N1} & \cdots & z_{1Np} \end{bmatrix}, \qquad N \times p,$$

(13)
$$Z_2 \equiv \begin{bmatrix} z_{211} & \cdots & z_{21q} \\ \vdots & & \vdots \\ z_{2T1} & , \ldots , & z_{2Tq} \end{bmatrix}, \qquad T \times q;$$

then (2) can be rewritten as:

(14) $y = X\beta + (I_T \otimes Z_1)\alpha_1 + (Z_2 \otimes I_N)\alpha_2 + u.$

Next, we list our assumptions with respect to the random vectors. α_1, α_2. Let, for arbitrary m, ι_m be an m-vector of unit elements, and define:

(15) $\bar{\alpha}_1 \equiv (\bar{\alpha}_{11}, \ldots, \bar{\alpha}_{1p})'$, $p \times 1$,

(16) $\bar{\alpha}_2 \equiv (\bar{\alpha}_{21}, \ldots, \bar{\alpha}_{2q})'$, $q \times 1$,

with $\bar{\alpha}_{1i}$ $(i = 1, \ldots, p)$ and $\bar{\alpha}_{2j}$ $(j = 1, \ldots, q)$ constants; then we make the following assumptions on the first and second moments of the distribution of α_1 and α_2:

(17) $E(\alpha_1) = \iota_T \otimes \bar{\alpha}_1$, $Tp \times 1$,

(18) $E(\alpha_2) = \bar{\alpha}_2 \otimes \iota_N$, $Nq \times 1$,

(19) $V(\alpha_1) = I_T \otimes \Sigma_1$, $Tp \times Tp$,

(20) $V(\alpha_2) = \Sigma_2 \otimes I_N$, $Nq \times Nq$,

with E the expectation and V the variance operator, and with $\Sigma_1(p \times p)$ and $\Sigma_2(q \times q)$ matrices of usually unknown parameters. We assume moreover:

(21) $$\operatorname{cov}(\alpha_1, \alpha_2) = 0$$

and we assume that α_1 and α_2 are distributed independently of X, Z_1, Z_2, and u. Rewrite (14):

(22) $$y = X\beta + (\iota_T \otimes Z_1)\bar{\alpha}_1 + (Z_2 \otimes \iota_N)\bar{\alpha}_2 + \epsilon \equiv \tilde{X}\tilde{\beta} + \epsilon,$$

with

(23) $$\epsilon \equiv u + (I_T \otimes Z_1)(\alpha_1 - \iota_T \otimes \bar{\alpha}_1) + (Z_2 \otimes I_N)(\alpha_2 - \bar{\alpha}_2 \otimes \iota_N).$$

Let $\sigma^2 \equiv E(u_{nt}^2)$; then the above assumptions imply:

(24) $$\Omega \equiv E(\epsilon\epsilon') = \sigma^2 I_{NT} + I_T \otimes Z_1 \Sigma_1 Z_1' + Z_2 \Sigma_2 Z_2' \otimes I_N.$$

When the first columns of Z_1 and Z_2 are unit vectors, $\bar{\alpha}_{11}$ and $\bar{\alpha}_{21}$ are not separately identifiable (even if there is no unit vector in X). For simplicity of exposition we assume that such an identification problem is solved by deletion of the appropriate columns of $\tilde{X}$ and the corresponding elements of $\tilde{\beta}$.

The GLS-estimator of $\tilde{\beta}$ is

(25) $$\tilde{b} \equiv (\tilde{X}'\Omega^{-1}\tilde{X})^{-1}\tilde{X}'\Omega^{-1}y.$$

The application of this formula requires an estimate of Ω. In this paper we will not derive estimators for Ω, although some estimators will be discussed in Section 3.1. Various estimators have been derived in an earlier paper [20]. Our main purpose now is to provide a class of decompositions of Ω, which will be done in Section 2. Section 3 gives some interpretations and discusses the theoretical and practical importance of the decompositions. Section 4 concludes.

2. THE STRUCTURE OF Ω

Define the following four projection matrices:

(26) $$K_1 \equiv Z_1(Z_1'Z_1)^{-1}Z_1', \qquad M_1 \equiv I_N - K_1, \qquad N \times N,$$

(27) $$K_2 \equiv Z_2(Z_2'Z_2)^{-1}Z_2', \qquad M_2 \equiv I_T - K_2, \qquad T \times T.$$

Let Q_1 (of order $N \times (N-p)$) and $R_1(N \times p)$ be matrices with the following properties:

(28) $$Q_1 Q_1' = M_1, \qquad Q_1' Q_1 = I_{N-p},$$

(29) $$R_1 R_1' = K_1, \qquad R_1' R_1 = I_p.$$

Let $Q_2(T \times (T-q))$ and $R_2(T \times q)$ be matrices such that:

(30) $$Q_2 Q_2' = M_2, \qquad Q_2' Q_2 = I_{T-q},$$

(31) $$R_2 R_2' = K_2, \qquad R_2' R_2 = I_q.$$

The columns of Q_i and R_i $(i = 1,2)$ constitute orthonormal bases for the spaces spanned by the columns of M_i and K_i $(i = 1,2)$, respectively. Note that $Q_i' R_i = 0$, $i = 1,2$, i.e., the space spanned by Q_i is orthogonal to the space spanned by R_i, $i = 1,2$. The Q's and R's are not fully determined by (28)–(31). Next, define:

(32) $$V \equiv (Q_2 \otimes Q_1, Q_2 \otimes R_1, R_2 \otimes Q_1, R_2 \otimes R_1), \qquad NT \times NT.$$

In view of the properties of the Q's and R's, $V'V = VV' = I_{NT}$. So, V is orthonormal. Its columns (and rows) constitute an orthonormal basis of NT-dimensional space. Let:

(33) $$W_1 \equiv R_1' Z_1 \Sigma_1 Z_1' R_1, \qquad p \times p,$$

(34) $$W_2 \equiv R_2' Z_2 \Sigma_2 Z_2' R_1, \qquad q \times q.$$

So:

(35) $$R_i W_i R_i' = K_i Z_i \Sigma_i Z_i' K_i = Z_i \Sigma_i Z_i', \qquad i = 1,2.$$

With these definitions, we can rewrite Ω:

(36) $$\begin{aligned}\Omega &= \sigma^2 I_{NT} + I_T \otimes Z_1 \Sigma_1 Z_1' + Z_2 \Sigma_2 Z_2' \otimes I_N \\ &= \sigma^2 (M_2 \otimes M_1 + M_2 \otimes K_1 + K_2 \otimes M_1 + K_2 \otimes K_1) \\ &\quad + (M_2 \otimes Z_1 \Sigma_1 Z_1' + K_2 \otimes Z_1 \Sigma_1 Z_1') \\ &\quad + (Z_2 \Sigma_2 Z_2' \otimes M_1 + Z_2 \Sigma_2 \Sigma_2' \otimes K_1) \\ &= \sigma^2 (Q_2 Q_2' \otimes Q_1 Q_1' + Q_2 Q_2' \otimes R_1 R_1' \\ &\qquad + R_2 R_2' \otimes Q_1 Q_1' + R_2 R_2' \otimes R_1 R_1') \\ &\quad + (Q_2 Q_2' \otimes R_1 W_1 R_1' + R_2 R_2' \otimes R_1 W_1 R_1') \\ &\quad + (R_2 W_2 R_2' \otimes Q_1 Q_1' + R_2 W_2 R_2' \otimes R_1 R_1') \\ &= \sigma^2 (Q_2 Q_2' \otimes Q_1 Q_1') + Q_2 Q_2' \otimes R_1 (\sigma^2 I_p + W_1) R_1' \\ &\quad + R_2 (\sigma^2 I_q + W_2) R_2' \otimes Q_1 Q_1' \\ &\quad + (R_2 \otimes R_1)(\sigma^2 I_{pq} + I_q \otimes W_1 + W_2 \otimes I_p)(R_2 \otimes R_1)' \\ &= V \Psi V',\end{aligned}$$

where $\Psi(NT \times NT)$ is:

(37) $$\Psi = \begin{bmatrix} \sigma^2 I_{(T-q)(N-p)} & 0 & 0 & 0 \\ 0 & I_{T-q} \otimes A_1 & 0 & 0 \\ 0 & 0 & A_2 \otimes I_{N-p} & 0 \\ 0 & 0 & 0 & A \end{bmatrix}$$

with:

(38) $$A_1 \equiv \sigma^2 I_p + W_1,$$

(39) $$A_2 \equiv \sigma^2 I_q + W_2,$$

(40) $$A \equiv \sigma^2 I_{pq} + I_q \otimes W_1 + W_2 \otimes I_p.$$

So, V serves to transform Ω into a block-diagonal matrix. The first block is a scalar matrix. The second one is block-diagonal, containing $(T - q)$ times the matrix A. The third block has a similar structure, but with rows and columns rearranged. The fourth block exhibits a structure somewhat like Ω itself. Since the Q's and R's are not fully determined by (28)–(31), (36) represents a whole class of decompositions.

Just like the number of linearly independent eigenvectors associated with an eigenvalue equals its multiplicity, we will sometimes find it convenient to speak of the multiplicities of the diagonal-blocks of Ψ. The multiplicity of σ^2 is $(T - q)(N - p)$. A_1 corresponds to the $(T - q)p$-dimensional space spanned by $Q_2 \otimes R_1$, so we say that the $p \times p$-matrix A_1 has multiplicity $T - q$. Analogously, A_2 has multiplicity $N - p$. The $pq \times pq$-matrix A has multiplicity one since it corresponds to the pq-dimensional space spanned by $R_2 \otimes R_1$. Using the relation $\Omega = V\Psi V'$, two goals can easily be attained. First, we can use the freedom left in specifying the Q's and R's to obtain the spectral decomposition of Ω. Second, we can use it to obtain simple expressions for arbitrary powers of Ω.

2.1 *The Spectral Decomposition of* Ω

When the Q's and R's are such that A_1, A_2, and A are diagonal, Ψ is diagonal, too. Then Ψ evidently contains the eigenvalues of Ω on its diagonal, and V is the corresponding matrix of eigenvectors.

From (38)–(40) we see that diagonality of A_1, A_2, and A is equivalent to diagonality of W_1 and W_2. In that case (35) represents the spectral decomposition of $Z_i\Sigma_i Z_i'$, $i = 1,2$; R_i contains the eigenvectors corresponding to the nonzero eigenvalues of $Z_i\Sigma_i Z_i'$. Denote by W_i^* the diagonal matrix of nonzero eigenvalues of $Z_i\Sigma_i Z_i'$ (i.e., the eigenvalues of $\Sigma_i Z_i' Z_i$), and let R_i^* be the corresponding matrix of eigenvectors; then:

(41) $$Z_i\Sigma_i Z_i' = R_i^* W_i^* R_i^{*\prime}, \qquad i = 1,2.$$

TABLE I
EIGENVALUES AND EIGENVECTORS OF Ω

Eigenvalue	Multiplicity	Corresponding Eigenvector(s)
σ^2	$(T-q)(N-p)$	$Q_2 \otimes Q_1$
$\sigma^2 + (W_1^*)_{ii}, i = 1, \ldots, p$	$T-q$	$Q_2 \otimes$ (ith column of R_1^*)
$\sigma^2 + (W_2^*)_{jj}, j = 1, \ldots, q$	$N-p$	(jth column of R_2^*) $\otimes Q_1$
$\sigma^2 + (W_1^*)_{ii} + (W_2^*)_{jj}, i = 1, \ldots, p; j = 1, \ldots, q$	1	(ith column of R_1^*) $\otimes$ (jth column of R_2^*)

The spectral decomposition of Ω is summarized in Table I. Notice that the multiplicity of the eigenvalues is just the multiplicity of the blocks of Ψ introduced above.

2.2 *Powers of* Ω

An admissible choice for R_i is:

(42) $$R_i^{**} = Z_i H_i,$$

with $H_i \equiv (Z_i'Z_i)^{-1/2}$, i = 1,2. Then:

(43) $$W_i = H_i Z_i' Z_i \Sigma_i Z_i' Z_i H_i = (Z_i'Z_i)^{1/2}\Sigma_i(Z_i'Z_i)^{1/2},$$

for $i = 1,2$. Denote the values of A_1, A_2, and A for this choice of the W's by A_1^{**}, A_2^{**}, and A^{**} (cf., (38)–(40)). Then, for arbitrary γ, powers of Ψ are:

(44) $$\Psi^\gamma = \begin{bmatrix} \sigma^{2\gamma} I_{(N-p)(T-q)} & 0 & 0 & 0 \\ 0 & I_{T-q} \otimes A_1^{**\gamma} & 0 & 0 \\ 0 & 0 & A_2^{**\gamma} \otimes I_{N-p} & 0 \\ 0 & 0 & 0 & A^{**\gamma} \end{bmatrix}.$$

Using $\Omega^\gamma = V\Psi^\gamma V'$, there holds:

(45) $$\Omega^\gamma = \sigma^{2\gamma} M_2 \otimes M_1 + M_2 \otimes Z_1 H_1 A_1^{**\gamma} H_1 Z_1' + Z_2 H_2 A_2^{**\gamma} H_2 Z_2' \otimes M_1 + (Z_2 H_2 \otimes Z_1 H_1) A^{**\gamma} (H_2 Z_2' \otimes H_1 Z_1').$$

For the case of special interest $\gamma = -1$, this expression can be simplified to obtain:

(46) $$\Omega^{-1} = \sigma^{-2} M_2 \otimes M_1 + M_2 \otimes Z_1(\sigma^2 Z_1'Z_1 + Z_1'Z_1\Sigma_1 Z_1'Z_1)^{-1} Z_1' + Z_2(\sigma^2 Z_2'Z_2 + Z_2'Z_2\Sigma_2 Z_2'Z_2)^{-1} Z_2' \otimes M_1 + (Z_2 \otimes Z_1)(\sigma^2 Z_2'Z_2 \otimes Z_1'Z_1 + Z_2'Z_2 \otimes Z_1'Z_1\Sigma_1 Z_1'Z_1 + Z_2'Z_2\Sigma_2 Z_2'Z_2 \otimes Z_1'Z_1)^{-1} (Z_2 \otimes Z_1)'.$$

Another case of interest is $\gamma = -\frac{1}{2}$; premultiplication of (22) by $\Omega^{-1/2}$ yields a model with a scalar variance-covariance matrix for which OLS is the best linear unbiased estimation technique.

3. DISCUSSION

It is illuminating to consider the decomposition (36) for the special case $p = q = 1$, $Z_1 = \iota_N$, $Z_2 = \iota_T$, $\bar{\alpha}_1 = \bar{\alpha}_2$ ($= 0$ if there is a unit column in X), viz. the error components model. Then (38)–(40) become

$$(47) \qquad A_1 = \sigma^2 + N\sigma_1^2,$$

$$(48) \qquad A_2 = \sigma^2 + T\sigma_2^2,$$

$$(49) \qquad A = \sigma^2 + N\sigma_1^2 + T\sigma_2^2,$$

where σ_1^2 is the variance of the individual component (μ_n in (1)) and σ_2^2 is the variance of the period component (ν_t in (1)). The matrix Ψ, cf. (37), is now diagonal. Hence, for the EC model, (36) reduces to the spectral decomposition of Ω. Furthermore, the eigenvectors do not depend on unknown parameters.

Obviously, decomposition (36) reduces to a spectral decomposition because $p = q = 1$; if $Z_1 \neq \iota_N$, $Z_2 \neq \iota_T$, (36) is still a spectral decomposition. Thus we can slightly generalize the error components model while retaining the spectral decomposition of its variance-covariance matrix. For p or q larger than one, the spectral decomposition generally involves eigenvectors that depend on unknown parameters,[2] cf. Table I, and the version (42)–(44) is to be preferred in estimation, as will be discussed briefly.

3.1 *Decomposition and Estimation*

To begin with, the decomposition suggests some simple transformations of the model, which greatly facilitates estimation. Recall the definition of V by (32). Premultiplying (22) by $Q_2' \otimes Q_1'$ yields a model with variance-covariance matrix $\sigma^2 I_{(T-q)(N-p)}$, i.e., the upper left hand block of Ψ. It can easily be verified that the estimators of β and σ^2 obtained by OLS in the transformed model are identical to the ones obtained by performing OLS in model (2) under the assumption that α_1 and α_2 are fixed. The estimators are a generalization of the least squares dummy variables estimator for the EC-model. They are asymptotically efficient [**18**, Chapter 5].

If one premultiplies (22) by $Q_2' \otimes R_1^{**\prime}$, one obtains

$$(50) \qquad (Q_2' \otimes R_1^{**\prime})y = (Q_2' \otimes R_1^{**\prime})X\beta + (Q_2'\iota_T \otimes R_1^{**\prime}Z_1)\bar{\alpha}_1 + (Q_2' \otimes R_1^{**\prime})\epsilon.$$

[2] As observed by one referee, also for p or q larger than one, there are some special cases where the spectral decomposition does not depend on unknown parameters. An obvious example would be Σ_1 and Σ_2 scalar matrices and Z_1 and Z_2 composed of orthogonal vectors. In that case W_1 and W_2, and hence A_1, A_2, and A, are diagonal, so that Ψ (cf. (37)) is diagonal and the decomposition (36) is the spectral decomposition.

The error term $(Q_2' \otimes R_1^{**\prime})\epsilon$ has variance-covariance matrix $I_{T-q} \otimes A_1^{**}$, i.e., the second diagonal block of Ψ. Let $\bar{\epsilon}$ be the vector of OLS-residuals of (50) and define $\bar{E}(p \times (T-q))$ by vec $\bar{E} \equiv \bar{\epsilon}$. Then a simple consistent estimator of A_1^{**} is $(T-q)^{-1}\bar{E}\bar{E}'$. Using (33) and (38) and the estimator of σ^2 obtained above, one can next derive an estimator for Σ_1. Of course, one can also iterate by using the estimate of A_1^{**} to form GLS estimates of β and $\bar{\alpha}_1$ and use the residuals of that GLS regression to update the estimate of A_1^{**}. Upon convergence this procedure produces an ML estimator of A_1^{**} in model (50). The resulting estimator of Σ_1 is asymptotically efficient in model (22). Actually, iteration until convergence is not required for asymptotic efficiency. One step (i.e., once updating the estimate of A_1^{**}) is sufficient. Estimates for Σ_2 can be generated analogously by premultiplying (22) by $R_2^{**} \otimes Q_1$.

One can also use the decomposition (36) in a slightly different way to obtain asymptotically efficient estimators of Σ_1 and Σ_2, which are also very simple to compute: Assume momentarily that the error-vector is observable. According to (36) we can rewrite its variance covariance matrix Ω as:

$$\Omega = \sigma^2(M_2 \otimes M_1) + M_2 \otimes R_1^{**}A_1^{**}R_1^{**\prime} + R_2^{**}A_2^{**}R_2^{**\prime} \otimes M_1 + (R_2^{**} \otimes R_1^{**})A(R_2^{**} \otimes R_1^{**})'. \tag{51}$$

Ignoring the fact that A_1^{**}, A_2^{**}, and A^{**} are functionally related, it is straightforward to derive best quadratic unbiased estimators of σ^2, A_1^{**}, A_2^{**}, and A^{**} [**18**, Sec. 5.3]. Using (38)–(40), one can in principle obtain estimators of Σ_1 and Σ_2 from these estimators. Obviously, however, Σ_1 and Σ_2 are overidentified. A possible solution is to ignore (40) and solve Σ_1 and Σ_2 from (38) and (39). Intuitively one would expect this procedure to be asymptotically efficient since the multiplicities of A_1^{**} and A_2^{**} increase with N or T, whereas the neglected matrix A^{**} has a constant multiplicity. Furthermore, the estimator of A^{**} is a matrix of rank one. This also suggests that the estimator of A^{**} better be ignored. Indeed it can be shown that solving Σ_1 and Σ_2 from (38) and (39), using the best quadratic estimators of A_1^{**} and A_2^{**}, provides asymptotically efficient estimators of Σ_1 and Σ_2 [**18**, Section 5.4].

As ϵ is not observable it has to be replaced by a predictor. It can be shown that if one takes the OLS-residuals of the analysis of covariance version of the model (i.e., (14) with α_1 and α_2 fixed), the resulting estimators of Σ_1 and Σ_2 are asymptotically efficient [**18**, Section 5.4]. The formulae for the estimators are:

$$\hat{\Sigma}_1 = \frac{1}{T-q}(Z_1'Z_1)^{-1}Z_1'\hat{U}M_2\hat{U}'Z_1(Z_1'Z_1)^{-1} - \hat{\sigma}^2(Z_1'Z_1)^{-1}, \tag{52}$$

$$\hat{\Sigma}_2 = \frac{1}{N-p}(Z_2'Z_2)^{-1}Z_2'\hat{U}'M_1\hat{U}Z_2(Z_2'Z_2)^{-1} - \hat{\sigma}^2(Z_2'Z_2)^{-1}, \tag{53}$$

where $\hat{U}(N \times T)$ is defined as vec$(\hat{U}) \equiv \hat{u}$, $\hat{u}$ being the vector of OLS residuals of the FE model (14). The estimators can be computed in $O(NT)$-time and can be given a simple interpretation [**18**, Sec. 5.4; **19**]. For the EC model $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$ reduce to what have been called Nerlove's estimators by Maddala and Mount [**7**, **12**]. In [**20**] we have considered two alternative estimators of σ^2, Σ_1, Σ_2, viz. ML and MINQUE. These estimators are also asymptotically efficient, but are more cumbersome to apply; ML requires iterative solution of a nonlinear system of equations, whereas the formula for MINQUE is quite complicated and involves some arbitrariness. We refer to [**20**] for a discussion.

Each of the estimators of the elements of Ω yields an asymptotically efficient estimator for β, when inserted in (25). A Monte Carlo study of various estimators by Wansbeek [**18**, Section 6.3] suggests that the analysis of covariance estimator of β (which obtains by transforming model (22) by premultiplication with the first part of V, cf. (32) and the first paragraph of this section) performs better than the GLS-methods based on MINQUE or based on the estimators (52) and (53).[3] In many cases, the latter estimator comes close however. ML often fails to converge. As an estimator of the elements of Ω both (52)–(53) and MINQUE show considerable variance, but the former estimator does better than MINQUE. These results suggest that the estimators based on the decomposition of Ω perform better than ML or MINQUE.

4. CONCLUSIONS

We have stressed the role of decompositions of the variance-covariance matrix of the errors in the analysis of linear models of panel data. From our generalization of the EC model it becomes obvious that a spectral decomposition is not necessarily the most convenient decomposition. The crucial feature of the decomposition involving A_1^{**}, A_2^{**}, A^{**} is that only Ψ depends on unknown parameters but not V. This allows for transformations of the model, before estimation, which suggest simple and efficient estimation methods of both the parameters of the systematic part and the elements of the variance-covariance matrix of the errors.

Netherlands Central Bureau of Statistics
and
University of Southern California

Manuscript received September, 1980; revision received June, 1981.

[3]The relatively good performance of the covariance estimator in these simulations contrasts with recent results obtained by Taylor [**16**] for an error components model with only individual effects. He proves that generally a feasible GLS estimator of β has lower variance than the covariance estimator. The simulations in [**18**] concern models with $k = p = q = 2$. Consequently, the variance-covariance matrix Ω has seven unknown parameters, as opposed to two in Taylor's case. Thus there is more room for error in estimating Ω and this may have a negative effect on the performance of the GLS estimator relative to the covariance estimator. Notice, incidentally, that Taylor's feasible GLS-estimator employs an estimate of the error variance-covariance matrix based on its spectral decomposition.

REFERENCES

[1] BALESTRA, P.: "Best Quadratic Unbiased Estimators of the Variance-Covariance Matrix in Normal Regression," *Journal of Econometrics*, 1(1973), 17–28.

[2] BALESTRA, P., AND M. NERLOVE: "Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas," *Econometrica*, 34(1966), 585–612.

[3] HILDRETH, C.: "Combining Cross Section-Data and Time Series," Cowles Commission Discussion Paper, Statistics No. 347, 1950.

[4] HSIAO, C.: "Statistical Inference for a Model with Both Random Cross-Sectional and Time Effects," *International Economic Review*, 15(1974), 12–30.

[5] ——— "Some Estimation Methods for a Random Coefficients Model," *Econometrica*, 43(1975), 305–326.

[6] KUH, E.: "The Validity of Cross-Sectionally Estimated Behavior Equations in Time Series Applications," *Econometrica*, 27(1959), 197–214.

[7] MADDALA, G. S. AND T. D. MOUNT: "A Comparative Study of Alternative Estimators for Variance Components Models Used in Econometric Applications," *Journal of the American Statistical Association*, 68(1973), 324–328.

[8] MALINVAUD, E., AND J. C. MILLERON: "Généralisation des modèles d'analyse de la variance dans le cas de séries temporelles de coupes instantanées. Conditions d'identifiabilité et méthodes d'estimation," INSEE Report 348/930, Paris, 1970.

[9] MAZODIER, P. (ED.): "The Econometrics of Panel Data," special issue of *Annales de l'INSEE*, 30–31(1978).

[10] MAZODIER, P., AND A. TROGNON: "Heteroscedasticity and Stratification in Error Components Models," *Annales de l'INSEE*, 30–31(1978), 451–482.

[11] MUNDLAK, Y.: "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46(1978), 69–86.

[12] NERLOVE, M.: "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections," *Econometrica*, 39(1971), 359–382.

[13] ———: "A Note on Error Components Models," *Econometrica*, 39(1971), 383–396.

[14] RAO, C. R.: "Estimation of Variance and Covariance Components—MINQUE Theory," *Journal of Multivariate Analysis*, 1(1971), 257–275.

[15] SWAMY, P. A. V. B., AND J. S. MEHTA: "Estimation of Linear Models with Time and Cross-Sectionally Varying Coefficients," *Journal of the American Statistical Association*, 72(1977), 890–898.

[16] TAYLOR, W. E.: "Small Sample Considerations in Estimation from Panel Data," *Journal of Econometrics*, 13(1980), 203–223.

[17] WALLACE, T. D., AND A. HUSSAIN: "The Use of Error Component Models in Combining Cross Section with Time Series Data," *Econometrica*, 37(1969), 55–68.

[18] WANSBEEK, T. J.: "Quantitative Effects in Panel Data Modelling," unpublished Ph.D. Thesis, Leyden University, 1980.

[19] WANSBEEK, T. J., AND A. KAPTEYN: "The Separation of Individual Variation and Systematic Change in the Analysis of Panel Survey Data," *Annales de l'INSEE*, 30–31(1978), 659–680.

[20] ———: "Estimators of the Covariance Structure of a Model for Longitudinal Data," in *Proceedings of the Econometric Society European Meeting 1979*, ed. by E. Charatsis. Amsterdam: North-Holland, 1981.