ON THE ESTIMATED VARIANCES OF REGRESSION COEFFICIENTS IN MISSPECIFIED ERROR COMPONENTS MODELS

PHILIPPE J. DESCHAMPS
University of Fribourg

In a regression model with an arbitrary number of error components, the co-variance matrix of the disturbances has three equivalent representations as linear combinations of matrices. Furthermore, this property is invariant with respect to powers, matrix addition, and matrix multiplication. This result is applied to the derivation and interpretation of the inconsistency of the estimated co-efficient variances when the error components structure is improperly restricted. This inconsistency is defined as the difference between the asymptotic variance obtained when the restricted model is correctly specified, and the asymptotic variance obtained when the restricted model is incorrectly specified; when some error components are improperly omitted, and the remaining variance compo-nents are consistently estimated, it is always negative. In the case where the time component is improperly omitted from the two-way model, we show that the difference between the true and estimated coefficient variances is of order greater than N^{-1} in probability.

1. INTRODUCTION

Since the pioneering article by Balestra and Nerlove [3], many authors have studied the estimation of linear regression models with error components. The Balestra–Nerlove specification was limited to a two-way classification of data, and omitted time-specific effects. This restriction was later lifted by Wallace and Hussain [10] and Nerlove [8], who considered the following er-ror specification:

$$u_{it} = v_i + w_t + \epsilon_{it} \qquad \text{for } i = 1, \ldots, n \quad \text{and} \quad t = 1, \ldots, T \tag{1}$$

where v_i, w_t, and ϵ_{it} have zero expectations, constant variances, and are pairwise and serially uncorrelated. Nerlove, in particular, found the spectral form of the error covariance matrix V implied by equation (1), expressing V as a linear combination of symmetric, idempotent, and pairwise orthog-

369

onal matrices that sum to an identity matrix. All the rational powers of V can be similarly represented; this is particularly useful for interpreting and computing the Aitken estimators of the regression coefficients. The spectral form also leads to a particularly simple expression of the ANOVA table for the variance components of u_{it} (for a classical presentation of this table, see, e.g., Graybill [6, p. 349]).

More recently, Searle and Henderson [9] and Deschamps [4] have studied the structure of the error covariance matrix in a model with a general, p-way classification of data (where there are $2^p - 1$ error components). As will be shown, their results imply that V has three equivalent representations as linear combinations of matrices. The coefficients of the first one are the variance components; the coefficients of the second one (the spectral form) are the eigenvalues of V; the coefficients of the third one are the elements of V. Each linear combination is, in a sense, isomorphic to the two others. Furthermore, the property is invariant with respect to powers, linear combinations, and matrix multiplication. This means that if the nonsingular matrices V_1, V_2, V_3, V_4 have the same form as V and if $a, b, \alpha, \beta, \gamma, \delta$ are scalars, then $V_1^\alpha (a V_2^\beta + b V_3^\gamma) V_4^\delta$ also has the same form; it can be represented as three linear combinations of the same matrices that appear in the three expressions for V. Furthermore, the three sets of coefficients are easily derived.

Section 2 of this paper will present the assumptions of the model, and a new notation. In Section 3, we present Theorem 1, which is a general result on the misspecified error components model. Theorem 1 can be used to derive and interpret the inconsistency (plim $N(\hat{\sigma}_i^2 - \sigma_i^2)$) of the estimated variances of the regression coefficients when some error components are improperly omitted (even though their variances are nonzero), and the remaining variance components are consistently estimated. We will show that the true coefficient variances are *always underestimated* in the misspecified model, for any p-way classification and for any number of omitted components. More generally, we find that the inconsistency can be expressed as three linear combinations of quadratic forms, in the same way as there are three representations of V. This turns out to be very helpful for the interpretation of the inconsistency.

When a *single* error component is omitted in the misspecified model and when the remaining variance components are consistently estimated, our results also imply that the inconsistency is *proportional* to the variance of the omitted component. The coefficient of proportionality is a linear combination of sums of squares of partial sums (as in, e.g., $\alpha_1 \sum_t (\sum_i Z_{it})^2 + \alpha_2 \sum_i (\sum_t Z_{it})^2$); it can be estimated as a by-product of generalized least squares on the misspecified model. This greatly facilitates the sensitivity analysis of misspecification.

Theorem 1 also applies to other types of misspecification: the inconsistency of an estimated coefficient variance is a linear combination of the inconsistencies of the variance component estimates.

In Section 4, we illustrate the previous results by analyzing the inconsistency when the time-specific component w_t is improperly omitted from equation (1). This case is of particular interest since it corresponds to the earliest (Balestra–Nerlove) specification, and is therefore found in most empirical implementations. In this instance the inconsistency has a particularly simple, intuitively appealing form, whose properties are easily investigated. It will be shown that this inconsistency is unbounded, unless the matrix of regressors satisfies very restrictive assumptions.

Lemma 1 of Appendix A presents the three isomorphic representations of V, and can be viewed as a synthesis of Searle and Henderson [9] and Deschamps [4]. This lemma will be used in Appendix B, where we prove Theorem 1 of Section 3. Appendix C presents estimators of the variance components which are consistent under misspecification, and coincide with the ones proposed by Amemiya [1] for a correctly specified three-component model.

2. A GENERAL ERROR COMPONENTS MODEL

Our regression model has 2^p error components, one of which is identically zero. It can be written as

$$y = X\beta + u \tag{2}$$

where y is an $N \times 1$ vector of observations on a dependent variable, X is an $N \times k$ matrix of observations on nonstochastic regressors, β is a $k \times 1$ vector of regression coefficients, and u is an $N \times 1$ vector of compound disturbances with the following structure:

$$u = \sum_{\alpha=00\ldots0}^{11\ldots1} (s_{n_1}^{1-\alpha_1} \otimes \cdots \otimes s_{n_p}^{1-\alpha_p}) v_\alpha \tag{3a}$$

with

$$E(v_\alpha) = 0_{q_\alpha} \tag{3b}$$

$$E(v_\alpha v_\alpha') = \theta_\alpha I_{q_\alpha} \tag{3c}$$

$$E(v_\alpha v_\beta') = O_{q_\alpha \times q_\beta} \qquad (\alpha \neq \beta). \tag{3d}$$

In equations (3a)–(3d) and in the rest of this paper, the Greek *subscripts* α, β, and γ indicate *binary* numbers with p digits, for example, $\alpha = \alpha_1 \alpha_2 \ldots \alpha_p$, where α_i is either zero or one. The letters $n_1, \ldots, n_p$ denote numbers of cells with $\prod_{i=1}^p n_i = N$, s_{n_i} is an $n_i \times 1$ vector of ones, and we adopt the convention that $s_{n_i}^0 = I_{n_i}$, an identity matrix of order n_i. v_α is the vector of the $q_\alpha = \prod_{i=1}^p n_i^{\alpha_i}$ realizations of error component α. This vector is assumed to have zero expectation and scalar covariance matrix in (3b) and (3c); equation (3d) assumes that the error components are pairwise uncorrelated. We take the *first* error component to be identically zero, and this implies that $\theta_{00\ldots0} = 0$.

Specification (3a)–(3d) obviously implies the following form for the covariance matrix of u (see Note 1):

$$V = E(uu') = \sum_{\alpha=00\ldots0}^{11\ldots1} \theta_\alpha (S_{n_1}^{1-\alpha_1} \otimes \cdots \otimes S_{n_p}^{1-\alpha_p}) \tag{4}$$

where S_{n_i} is a square matrix of ones of order n_i and $S_{n_i}^0 \equiv I_{n_i}$.

The model in Nerlove [8] is obtained as a special case of (4) by letting $p = 2$, $n_1 = n$, and $n_2 = T$:

$$V = \theta_{00} S_{nT} + \theta_{01}(S_n \otimes I_T) + \theta_{10}(I_n \otimes S_T) + \theta_{11} I_{nT} \tag{5}$$

where $\theta_{00} = 0$, $\theta_{01} \equiv \sigma_w^2$, $\theta_{10} \equiv \sigma_v^2$, and $\theta_{11} \equiv \sigma_\epsilon^2$ (compare equation (1)). By letting $\theta_{01} = 0$ in (5), the model specializes further to Balestra and Nerlove [3].

The reader will have no difficulty in verifying that when $p = 3$, if $\theta_{000} = \theta_{001} = \theta_{010} = \theta_{011} = \theta_{101} = 0$ and if $\theta_{100} \equiv \sigma_v^2$, $\theta_{110} \equiv \sigma_w^2$, $\theta_{111} \equiv \sigma_\epsilon^2$, equation (4) specializes to the nested specification in Fuller and Battese [5], implied by

$$u_{ijt} = v_i + w_{ij} + \epsilon_{ijt}. \tag{6}$$

So the binary number α may serve to interpret the variance component θ_α in equation (4): it is simply the variance of that error term which has the particular index subset identified by the unitary digits in α.

Equation (4) is the first of our three isomorphic characterizations of V; the two others are given in Appendix A.

3. THE CONSEQUENCES OF MISSPECIFICATION

We consider the case where V in equation (4) is estimated by:

$$\hat{V}_* = \sum_{\alpha=00\ldots0}^{11\ldots1} \hat{\theta}_\alpha^* \left(\bigotimes_{i=1}^p S_{n_i}^{1-\alpha_i} \right) \tag{7}$$

where $\hat{\theta}^*$ is a possibly inconsistent estimator of the $2^p \times 1$ vector θ. We estimate β by

$$\hat{\beta} = (X'\hat{V}_*^{-1}X)^{-1}X'\hat{V}_*^{-1}y. \tag{8}$$

The covariance matrix of $\hat{\beta}$ is estimated by:

$$C = (X'\hat{V}_*^{-1}X)^{-1}$$

whereas the true covariance matrix, conditional on $\hat{\theta}^*$, is

$$V(\hat{\beta}|\hat{\theta}^*) = (X'\hat{V}_*^{-1}X)^{-1}X'\hat{V}_*^{-1}V\hat{V}_*^{-1}X(X'\hat{V}_*^{-1}X)^{-1}.$$

We let $\theta^* = \text{plim } \hat{\theta}^*$ (see Note 2), and assume that $V_* = \sum_\alpha \theta_\alpha^* \bigotimes_{i=1}^p S_{n_i}^{1-\alpha_i}$ is positive definite. We also assume that $X'V_*^{-1}X/N$ tends to a finite, positive definite matrix M_*. We further note that

$$NC = \left(\frac{X'\hat{V}_*^{-1}X}{N} \right)^{-1}$$

and

$$NV(\hat{\beta}|\hat{\theta}^*) = \left(\frac{X'\hat{V}_*^{-1}X}{N}\right)^{-1}\left(\frac{X'\hat{V}_*^{-1}V\hat{V}_*^{-1}X}{N}\right)\left(\frac{X'\hat{V}_*^{-1}X}{N}\right)^{-1}$$

We wish to evaluate[3]

$\text{plim } N(C - V(\hat{\beta}|\hat{\theta}^*))$

$\quad = \text{plim } N((X'\hat{V}_*^{-1}X)^{-1} - (X'\hat{V}_*^{-1}X)^{-1}X'\hat{V}_*^{-1}V\hat{V}_*^{-1}X(X'\hat{V}_*^{-1}X)^{-1})$

$\quad = \lim N((X'V_*^{-1}X)^{-1} - (X'V_*^{-1}X)^{-1}X'V_*^{-1}VV_*^{-1}X(X'V_*^{-1}X)^{-1})$

$\quad = \lim N((X'V_*^{-1}X)^{-1}(X'V_*^{-1}V_*V_*^{-1}X - X'V_*^{-1}VV_*^{-1}X)(X'V_*^{-1}X)^{-1})$

$$\quad = M_*^{-1} \lim \left(\frac{X'_* D X_*}{N}\right) M_*^{-1} \tag{9}$$

with $X_* = V_*^{-1/2}X$ and $D = V_*^{-1/2}(V_* - V)V_*^{-1/2}$.

Equation (9) implies the following expression for the inconsistency of the estimated variance of $\sqrt{N}\hat{\beta}_\ell$:

$$\text{plim } N(C_{\ell\ell} - V(\hat{\beta}_\ell|\hat{\theta}^*)) = \lim \left(\frac{Z'(\ell)DZ(\ell)}{N}\right) \tag{10}$$

where $Z(\ell)$ is the ℓth column of the matrix $X_* M_*^{-1}$.

The following definition provides the key to the three equivalent representations of V given in Appendix A, and to the three corresponding representations of (10) given by Theorem 1 of this section.

DEFINITION 1. *Let, as before, S_{n_i} be a matrix of ones of order n_i and zero powers denote identity matrices. We define*

$$K = \bigotimes_{i=1}^{p} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}; \tag{11}$$

$$K_0 = \bigotimes_{i=1}^{p} \begin{pmatrix} n_i & 1 \\ 0 & 1 \end{pmatrix}; \tag{12}$$

$$L_\beta = \bigotimes_{i=1}^{p} S_{n_i}^{1-\beta_i}; \tag{13}$$

$$R_\beta = \bigotimes_{i=1}^{p} \left(\frac{1}{n_i} S_{n_i}\right)^{1-\beta_i} \left(I_{n_i} - \frac{1}{n_i} S_{n_i}\right)^{\beta_i}; \tag{14}$$

$$M_\beta = \bigotimes_{i=1}^{p} (S_{n_i} - I_{n_i})^{1-\beta_i}. \tag{15} \quad \blacksquare$$

We now define $\eta = \theta^* - \theta$ as the vector of variance component inconsistencies. It is shown in Appendix A that the elements of $K_0\theta^*$ are the eigenvalues of V_*, so that they are strictly positive by assumption. Theorem 1 shows that the inconsistency (10) of an estimated coefficient variance is a

linear combination of the elements of η, with coefficients that depend on θ^*, but not on θ.

THEOREM 1. *Let* $\eta = \text{plim } \hat{\theta}^* - \theta = \theta^* - \theta$, *let* $V_* = \sum_\alpha \theta_\alpha^* L_\alpha$, *let* $D = V_*^{-1/2}(V_* - V)V_*^{-1/2}$, *and let* Δ *be the diagonal matrix with the elements of* $K_0\theta^*$ *on the diagonal. Then* D *can be equivalently expressed as*

$$D = \sum_{\gamma=00\ldots0}^{11\ldots1} \eta_\gamma \sum_{\beta=00\ldots0}^{11\ldots1} a_{\beta\gamma} R_\beta \qquad (16)$$

$$D = \sum_{\gamma=00\ldots0}^{11\ldots1} \eta_\gamma \sum_{\beta=00\ldots0}^{11\ldots1} b_{\beta\gamma} L_\beta \qquad (17)$$

$$D = \sum_{\gamma=00\ldots0}^{11\ldots1} \eta_\gamma \sum_{\beta=00\ldots0}^{11\ldots1} c_{\beta\gamma} M_\beta \qquad (18)$$

where $a_{\beta\gamma}$, $b_{\beta\gamma}$, *and* $c_{\beta\gamma}$ *are the elements in row* β *and column* γ *of* $\Delta^{-1}K_0$, $K_0^{-1}\Delta^{-1}K_0$, *and* $KK_0^{-1}\Delta^{-1}K_0$, *respectively.*   ∎

The most interesting type of misspecification occurs when one, or several, variance components θ_γ are improperly excluded from equation (4). This follows from letting $\hat{\theta}_\gamma^* = 0$, so that $\theta_\gamma^* = 0$. It is shown in Appendix C that the remaining variance components can always be consistently estimated. In this case the inconsistency (10) is always negative: indeed, it is shown in Appendix A that the matrices R_β in (16) are idempotent, pairwise orthogonal, and add up to an identity matrix, so that the eigenvalues of D are $\lambda_\beta = \sum_\gamma \eta_\gamma a_{\beta\gamma}$. Since $a_{\beta\gamma} \geq 0$ and since $\eta_\gamma = 0$ when component γ is included, whereas $\eta_\gamma = -\theta_\gamma \leq 0$ when component γ is excluded, we have that $\lambda_\beta \leq 0$ for all β.

We may also note that the opposite type of misspecification whereby a variance component θ_γ is improperly *included* is of no consequence asymptotically if plim $\hat{\theta}_\gamma^* = \theta_\gamma = 0$, so that $\eta_\gamma = 0$ in Theorem 1.

Equation (17) implies that the quadratic form in (10) is a linear combination of sums of squares of partial sums of the elements of $Z(\ell)$, as is obvious from the definition of L_β in (13). This fact will be used in Section 4. Similarly, equation (18) expresses the quadratic form as a linear combination of sums of cross-products of the elements of $Z(\ell)$ (see the definition of M_β in (15)). Using the arguments in Appendix A, it is easy to show that the coefficients $\sum_\gamma \eta_\gamma c_{\beta\gamma}$ in equation (18) are in fact the elements of D.

4. AN ILLUSTRATION

In this section, we illustrate the preceding theorem by considering the case where w_t is improperly omitted from equation (1); equivalently, $\hat{\theta}_{01}^* = 0$ in equation (7). We will first consider the case where the two remaining variance components are consistently estimated (Appendix C shows that this will be true if the two-way formulas are used). This assumption will be relaxed

at the end of this section. For now, we assume that $\theta_{00}^* = \theta_{01}^* = 0$, $\theta_{10}^* = \theta_{10}$, and $\theta_{11}^* = \theta_{11}$.

As previously noted, the elements of $K_0\theta^*$ in Theorem 1 are the eigenvalues of V_*. Using equation (12), they may be written as

$$
\begin{bmatrix} \lambda_{00} \\ \lambda_{01} \\ \lambda_{10} \\ \lambda_{11} \end{bmatrix} = \begin{bmatrix} nT & n & T & 1 \\ 0 & n & 0 & 1 \\ 0 & 0 & T & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \theta_{10} \\ \theta_{11} \end{bmatrix} = \begin{bmatrix} T\theta_{10} + \theta_{11} \\ \theta_{11} \\ T\theta_{10} + \theta_{11} \\ \theta_{11} \end{bmatrix}. \tag{19}
$$

Furthermore, the coefficients $b_{\beta\gamma}$ in equation (17) are elements of the matrix:

$$K_0^{-1}\Delta^{-1}K_0$$

$$
= \frac{1}{nT} \begin{bmatrix} 1 & -1 & -1 & 1 \\ 0 & T & 0 & -T \\ 0 & 0 & n & -n \\ 0 & 0 & 0 & nT \end{bmatrix} \begin{bmatrix} \lambda_{00}^{-1} & 0 & 0 & 0 \\ 0 & \lambda_{01}^{-1} & 0 & 0 \\ 0 & 0 & \lambda_{10}^{-1} & 0 \\ 0 & 0 & 0 & \lambda_{11}^{-1} \end{bmatrix} \begin{bmatrix} nT & n & T & 1 \\ 0 & n & 0 & 1 \\ 0 & 0 & T & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$
= \frac{1}{nT} \begin{bmatrix} nT\lambda_{00}^{-1} & n(\lambda_{00}^{-1} - \lambda_{01}^{-1}) & T(\lambda_{00}^{-1} - \lambda_{10}^{-1}) & \lambda_{00}^{-1} - \lambda_{01}^{-1} - \lambda_{10}^{-1} + \lambda_{11}^{-1} \\ 0 & nT\lambda_{01}^{-1} & 0 & T(\lambda_{01}^{-1} - \lambda_{11}^{-1}) \\ 0 & 0 & nT\lambda_{10}^{-1} & n(\lambda_{10}^{-1} - \lambda_{11}^{-1}) \\ 0 & 0 & 0 & nT\lambda_{11}^{-1} \end{bmatrix}. \tag{20}
$$

Since the only inconsistency occurs in the time component, we have $\eta_{01} = -\theta_{01}$ and $\eta_{00} = \eta_{10} = \eta_{11} = 0$, so that we only need those coefficients $b_{\beta\gamma}$ that are located in the second column of the preceding matrix (which has the binary index 01). Upon substituting (13), (19), and (20) in equation (17), we obtain

$$
D = -\theta_{01}\left(\frac{1}{T}\left(\frac{1}{T\theta_{10} + \theta_{11}} - \frac{1}{\theta_{11}} \right) S_{nT} + \frac{1}{\theta_{11}}(S_n \otimes I_T) \right)
$$

$$
= -\frac{\theta_{01}}{\theta_{11}}\left(\frac{-\theta_{10}}{T\theta_{10} + \theta_{11}} S_{nT} + (S_n \otimes I_T) \right) \tag{21}
$$

so that the quadratic form in equation (10) is

$$
Z'(\ell)DZ(\ell) = -\frac{\theta_{01}}{\theta_{11}}\left(\sum_{t=1}^{T} Z_{.t}^2(\ell) - \frac{\theta_{10}}{T\theta_{10} + \theta_{11}} Z_{..}^2(\ell) \right) \tag{22}
$$

with $Z_{.t}(\ell) = \sum_{i=1}^{n} Z_{it}(\ell)$ and $Z_{..}(\ell) = \sum_{i=1}^{n}\sum_{t=1}^{T} Z_{it}(\ell)$.

Since $\theta_{10}/(T\theta_{10} + \theta_{11}) < 1/T$, an upper bound for the expression in (22) is given by

$$Z'(\ell)DZ(\ell) \le -\frac{\theta_{01}}{\theta_{11}} \sum_{t=1}^{T} \left(Z_{.t}(\ell) - \frac{Z_{..}(\ell)}{T} \right)^2$$

and it is seen that for given θ_{01} and θ_{11}, $\lim_{T\to\infty} Z'(\ell)DZ(\ell) = -\infty$, unless $\theta_{01} = 0$ or unless $Z_{.t}(\ell) = \alpha$ for all t. In the latter case, equation (22) is easily shown to imply

$$Z'(\ell)DZ(\ell) = \frac{-T\alpha^2\theta_{01}}{T\theta_{10} + \theta_{11}}$$

whose limit equals $-\alpha^2\theta_{01}/\theta_{10}$ as T tends to infinity.

To summarize, the inconsistency of the ℓth estimated variance in the misspecified model is unbounded as $n \to \infty$ and $T \to \infty$. This may be seen by dividing equation (22) by nT, and noting that $\sum_{t=1}^{T} Z_{.t}^2(\ell)/nT$ does not converge in general. The unboundedness occurs unless the "observations" in the ℓth column of $X_* M_*^{-1}$ repeat their average pattern over time. In this (very special) case where $Z_{.t}(\ell) = \alpha$ for all t, we have

$$\text{plim } nT(C_{\ell\ell} - V(\hat{\beta}_\ell|\hat{\theta}^*)) = \lim \frac{Z'(\ell)DZ(\ell)}{nT} = \lim \frac{1}{nT}\left(\frac{-\alpha^2\theta_{01}}{\theta_{10}} \right) = 0.$$

We will now show that the previous result remains essentially unchanged when $\hat{\theta}_{10}^*$ and $\hat{\theta}_{11}^*$ are the one-way estimators:

$$\hat{\theta}_{11}^* = \frac{1}{n(T-1)} (\hat{u}'(I_n \otimes R_1)\hat{u})$$

$$\hat{\theta}_{10}^* = \frac{1}{T}\left(\frac{\hat{u}'(I_n \otimes R_0)\hat{u}}{n} - \frac{\hat{u}'(I_n \otimes R_1)\hat{u}}{n(T-1)} \right)$$

where $R_0 = T^{-1}S_T$, $R_1 = I_T - R_0$, and where $\hat{u}$ is a predictor of u. If the true disturbances u_{it} are known and $\hat{u} = u$, it is easy to show that $E(\hat{\theta}_{11}^*) = \theta_{01} + \theta_{11}$ and $E(\hat{\theta}_{10}^*) = \theta_{10}$; we will therefore assume that $\theta_{11}^* = \theta_{01} + \theta_{11}$ and that $\theta_{10}^* = \theta_{10}$. As before, we have $\theta_{00}^* = \theta_{01}^* = 0$. The eigenvalues of V_* are again given by $K_0\theta^*$, this time as

$$\begin{pmatrix} \lambda_{00} \\ \lambda_{01} \\ \lambda_{10} \\ \lambda_{11} \end{pmatrix} = \begin{pmatrix} T\theta_{10} + \theta_{01} + \theta_{11} \\ \theta_{01} + \theta_{11} \\ T\theta_{10} + \theta_{01} + \theta_{11} \\ \theta_{01} + \theta_{11} \end{pmatrix}.$$

The vector η of variance component inconsistencies is this time given by:

$$
\begin{pmatrix} \eta_{00} \\ \eta_{01} \\ \eta_{10} \\ \eta_{11} \end{pmatrix} = \begin{pmatrix} 0 \\ -\theta_{01} \\ 0 \\ \theta_{01} \end{pmatrix}.
$$

Upon using again the methodology of Theorem 1, we see that $D = D_1 + D_2$, with

$$
D_1 = -\theta_{01}\left(\frac{1}{T}\left(\frac{1}{T\theta_{10} + \theta_{01} + \theta_{11}} - \frac{1}{\theta_{01} + \theta_{11}}\right)S_{nT} + \frac{1}{\theta_{01} + \theta_{11}}(S_n \otimes I_T)\right)
$$

$$
D_2 = \theta_{01}\left(\frac{1}{T}\left(\frac{1}{T\theta_{10} + \theta_{01} + \theta_{11}} - \frac{1}{\theta_{01} + \theta_{11}}\right)(I_n \otimes S_T) + \frac{1}{\theta_{01} + \theta_{11}}I_{nT}\right).
$$

The matrix D_1 has the same form as D in (21), with θ_{11} replaced by $\theta_{01} + \theta_{11}$; hence the preceding analysis also applies to $Z'(\ell)D_1 Z(\ell)/nT$, which is negative and generally unbounded. By analogy with (22), the second term may be written as

$$
Z'(\ell)D_2 Z(\ell) = \frac{\theta_{01}}{\theta_{01} + \theta_{11}}\left(\sum_{i=1}^{n}\sum_{t=1}^{T} Z_{it}^2(\ell) - \frac{\theta_{10}}{T\theta_{10} + \theta_{01} + \theta_{11}}\sum_{i=1}^{n} Z_{i.}^2(\ell)\right)
$$

$$
\geq \frac{\theta_{01}}{\theta_{01} + \theta_{11}}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(Z_{it}(\ell) - \frac{Z_{i.}(\ell)}{T}\right)^2 \geq 0 \tag{23}
$$

and $Z'(\ell)D_2 Z(\ell)/nT$ converges to a finite, nonnegative number if $\sum_{i,t} Z_{it}^2(\ell)/nT$ and $\sum_i Z_{i.}^2(\ell)/nT^2$ are bounded. Furthermore the lower bound in (23) vanishes if, and only if, $Z_{it}(\ell) = c_i$ for all t; in this case the limit of $Z'(\ell)D_2 Z(\ell)/nT$ is easily shown to vanish when $\theta_{10} \neq 0$.

Collecting results, we may say that the inconsistency remains negative and generally unbounded, since the negative term dominates the positive one. Furthermore, when $\theta_{10} \neq 0$, the inconsistency now vanishes if (and only if) the elements of the relevant column of $X_* M_*^{-1}$ *exactly* repeat their pattern over time.

5. CONCLUDING REMARKS

In a general error components model with a p-way classification of data, we have given three interpretations of the inconsistency of the estimated coefficient variances when the error components structure is misspecified. In the special case where the time component is improperly omitted from the two-

way classification model, this inconsistency is in general unbounded, in the sense that $C_{\ell\ell} - V(\hat{\beta}_\ell | \hat{\theta}^*)$ is of order greater than N^{-1} in probability, so that plim $NV(\hat{\beta}_\ell | \hat{\theta}^*)$ does not exist. Since plim $NV(\hat{\beta}_\ell | \hat{\theta}^*)$ is the asymptotic variance when the restricted model is incorrectly specified, this implies that $\sqrt{N}(\hat{\beta}_\ell - \beta_\ell)$ does not have a proper limiting distribution in such a case.

The arguments in Appendixes A and C indicate that the p-way model is essentially a tensor generalization of the 2-way model, and is, as such, not much more difficult to handle. It is therefore unfortunate, in view of the potentially serious consequences of misspecification, that most contributions to the literature on error components have been limited to the Balestra–Nerlove specification.

NOTES

1. Note that equation (4) differs from Searle and Henderson [9, equation 2.2] in the indexing of the variance components θ_α. A correspondence between the two indexing schemes is obtained by writing the digits in α in reverse order, and taking the complement to unity of each digit: for instance, θ_{011} in our notation corresponds to θ_{001} in [9]. As we will see, our notation has definite advantages for interpreting the variance components, and for stating the isomorphisms of Appendix A.

2. Unless otherwise indicated, all the limits in this paper are taken as $n_i \rightarrow \infty$ for all i.

3. If $\hat{V}_* = \hat{\sigma}_u^2 I$ with $\hat{\sigma}_u^2$ the OLS estimator of the error variance, it may be shown (Greenwald, [7]) that $(X'X)^{-1}X'(V_* - V)X(X'X)^{-1}$ approximates the *bias* of the estimated covariance matrix when N is large.

4. Note that $c_{\alpha\beta} = 0$ for $\beta = 00\ldots0$.

REFERENCES

1. Amemiya, T. The estimation of the variances in a variance components model. *International Economic Review* 12 (1971): 1–13.
2. Balestra, P. Best quadratic unbiased estimators of the variance-covariance matrix in normal regression. *Journal of Econometrics* 1 (1973): 17–28.
3. Balestra, P. & 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.
4. Deschamps, P. A note on isomorphic characterizations of the dispersion matrix in error components models. *Linear Algebra and Its Applications* 111 (1988): 147–150.
5. Fuller, W.A. & G.E. Battese. Transformations for estimation of linear models with nested-error structure. *Journal of the American Statistical Association* 68 (1973): 626–632.
6. Graybill, F.A. *An Introduction to Linear Statistical Models*, vol. I. New York: McGraw-Hill, 1961.
7. Greenwald, B.C. A general analysis of bias in the estimated standard errors of least squares coefficients. *Journal of Econometrics* 22 (1983): 323–338.
8. Nerlove, M. A note on error components models. *Econometrica* 39 (1971): 383–396.
9. Searle, S.R. & H.V. Henderson. Dispersion matrices for variance components models. *Journal of the American Statistical Association* 74 (1979): 465–470.
10. Wallace, T.D. & A. Hussain. The use of error components models in combining cross section with time series data. *Econometrica* 37 (1969): 55–72.

APPENDIX A. CHARACTERIZATIONS OF THE ERROR COVARIANCE MATRIX

It is easily shown that the following lemma can be extended to all the rational powers of V, and also to linear combinations or products of matrices having the same structure as V (an example of such an extension is in fact given by Theorem 1).

LEMMA 1. *Let u have the specification* $(3a)$–$(3d)$, *and let* $V = E(uu')$. *Then*

$$V = \sum_{\alpha=00\ldots0}^{11\ldots1} \theta_\alpha L_\alpha = \sum_{\alpha=00\ldots0}^{11\ldots1} \lambda_\alpha R_\alpha = \sum_{\alpha=00\ldots0}^{11\ldots1} \rho_\alpha M_\alpha$$

where L_α, R_α, *and* M_α *are given by Definition 1. The* λ_α *are the eigenvalues of V and the* ρ_α *are the elements of V. Furthermore, if* λ, ρ *and* θ *denote the* $2^p \times 1$ *vectors with elements* λ_α, ρ_α, *and* θ_α, *we have* $\lambda = K_0\theta$ *and* $\rho = K\theta$, *with K and* K_0 *given by* (11) *and* (12).

Proof. Equation (4) may also be written as

$$V = (\theta' \otimes I_N)L \tag{24}$$

where

$$L = \begin{pmatrix} L_{00\ldots0} \\ \vdots \\ L_{11\ldots1} \end{pmatrix}$$

and where the matrices L_α are as defined in (13). Using $\lambda = K_0\theta$ and $\rho = K\theta$, (24) is equivalent to

$$V = (\theta' \otimes I_N)(K_0' \otimes I_N)((K_0')^{-1} \otimes I_N)L$$
$$= (\lambda' \otimes I_N)((K_0')^{-1} \otimes I_N)L \equiv (\lambda' \otimes I_N)R \tag{25}$$

and

$$V = (\theta' \otimes I_N)(K' \otimes I_N)((K')^{-1} \otimes I_N)L$$
$$= (\rho' \otimes I_N)((K')^{-1} \otimes I_N)L \equiv (\rho' \otimes I_N)M. \tag{26}$$

If we partition R and M in the same fashion as L,

$$R = \begin{pmatrix} R_{00\ldots0} \\ \vdots \\ R_{11\ldots1} \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} M_{00\ldots0} \\ \vdots \\ M_{11\ldots1} \end{pmatrix},$$

it follows from (25) and (26) that

$$R_\beta = \sum_{\alpha=00\ldots0}^{11\ldots1} k_{\beta\alpha}^0 L_\alpha \tag{27}$$

and

$$M_\beta = \sum_{\alpha=00\ldots0}^{11\ldots1} k_{\beta\alpha} L_\alpha \tag{28}$$

where $k_{\beta\alpha}^0$ and $k_{\beta\alpha}$ are the elements in row β and column α of $(K_0')^{-1}$ and $(K')^{-1}$, respectively.

We now show that (27) is the expansion of (14). When $p = 1$, we have

$$(K_0')^{-1} = n_p^{-1}\begin{pmatrix} 1 & 0 \\ -1 & n_p \end{pmatrix}, \qquad L_0 = S_{n_p}, \qquad L_1 = I_{n_p},$$

so that (27) obviously implies (14). It is easy to see that the property is true for p whenever it is true for $p - 1$. Indeed, $(K_0')^{-1} \equiv (K_0'(p))^{-1}$ and $L \equiv L(p)$ may be defined recursively as

$$(K_0'(p))^{-1} = \begin{pmatrix} n_1^{-1}(K_0'(p-1))^{-1} & O \\ -n_1^{-1}(K_0'(p-1))^{-1} & (K_0'(p-1))^{-1} \end{pmatrix}$$

$$L(p) = \begin{bmatrix} S_{n_1} \otimes L_{00\ldots0}(p-1) \\ \vdots \\ S_{n_1} \otimes L_{11\ldots1}(p-1) \\ I_{n_1} \otimes L_{00\ldots0}(p-1) \\ \vdots \\ I_{n_1} \otimes L_{11\ldots1}(p-1) \end{bmatrix}$$

so that, upon letting $\beta = (\beta_1 \gamma_1 \gamma_2 \cdots \gamma_{p-1})$, (27) may be written if $\beta_1 = 0$ as

$$R_\beta(p) = \sum_\alpha (n_1^{-1} k_{\gamma\alpha}^0(p-1))(S_{n_1} \otimes L_\alpha(p-1))$$

$$= n_1^{-1} S_{n_1} \otimes \sum_\alpha k_{\gamma\alpha}^0(p-1) L_\alpha(p-1)$$

$$= n_1^{-1} S_{n_1} \otimes R_\gamma(p-1)$$

and if $\beta_1 = 1$ as:

$$R_\beta(p) = \sum_\alpha (-n_1^{-1} k_{\gamma\alpha}^0(p-1))(S_{n_1} \otimes L_\alpha(p-1))$$

$$+ \sum_\alpha (k_{\gamma\alpha}^0(p-1))(I_{n_1} \otimes L_\alpha(p-1))$$

$$= (I_{n_1} - n_1^{-1} S_{n_1}) \otimes \sum_\alpha k_{\gamma\alpha}^0(p-1) L_\alpha(p-1)$$

$$= (I_{n_1} - n_1^{-1} S_{n_1}) \otimes R_\gamma(p-1)$$

where we note that α and γ now involve $p - 1$ binary digits rather than p. The proof that (28) is the expansion of (15) is exactly similar.

We must show that the matrices R_β in (14) are idempotent, pairwise orthogonal, and add up to an identity matrix, thus proving our claim that λ is indeed the vector

of eigenvalues. $R_\beta R_\beta = R_\beta$ is obvious since R_β is a Kronecker product of idempotent matrices. $R_\alpha R_\beta = O$ for $\alpha \neq \beta$ follows from the fact that $n_i^{-1} S_{n_i}(I_{n_i} - n_i^{-1} S_{n_i}) = O$; indeed, $\alpha \neq \beta$ implies the existence of an index i such that $\alpha_i = 0$ and $\beta_i = 1$, or $\alpha_i = 1$ and $\beta_i = 0$. So $\sum_\beta R_\beta$ is symmetric and idempotent; in order to show that it equals I_N, we show that its trace is nonzero. This follows easily from (27), which implies

$$\text{tr}\left(\sum_{\beta=00\ldots0}^{11\ldots1} R_\beta\right) = \sum_{\beta=00\ldots0}^{11\ldots1} \sum_{\alpha=00\ldots0}^{11\ldots1} k_{\beta\alpha}^0 \, \text{tr}(L_\alpha) = N \sum_{\beta=00\ldots0}^{11\ldots1} \sum_{\alpha=00\ldots0}^{11\ldots1} k_{\beta\alpha}^0 = N$$

since $\sum_\beta \sum_\alpha k_{\beta\alpha}^0$ is the sum of the elements of $(K_0')^{-1}$, which is easily seen to equal unity:

$$s'(K_0')^{-1}s = \left(\bigotimes_{i=1}^{p} (1\ \ 1)\right)\left(\bigotimes_{i=1}^{p}\begin{pmatrix} n_i^{-1} & 0 \\ -n_i^{-1} & 1 \end{pmatrix}\right)\left(\bigotimes_{i=1}^{p}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = 1.$$

It was shown in Deschamps [4] that the elements of ρ are also the elements of V. Since, from (15), the elements of M_β are 0 or 1, this is also implied by $\sum_\beta M_\beta = S_N$, which is readily verified from (28).

APPENDIX B. PROOF OF THEOREM 1

It is an immediate consequence of Lemma 1 that V, V_*, and $V_* - V$ can be expressed as (different) linear combinations of the same matrices R_α. The coefficients in the linear combinations are the eigenvalues; in the case of $V_* - V$, they form the vector $K_0\theta^* - K_0\theta = K_0\eta$.

As is easily seen from the properties of the matrices R_α, both $V_*^{-1/2}$ and D can again be represented as linear combinations of the same matrices R_α. The eigenvalues of D are obtained by multiplying the eigenvalues of $V_* - V$ by those of V_*^{-1}, so that they form the vector $\Delta^{-1}K_0\eta$, with elements $\sum_\gamma \eta_\gamma a_{\beta\gamma}$. This proves that (16) is indeed the spectral form of D.

In order to obtain equation (17), we substitute (27) into (16):

$$D = \sum_{\gamma=00\ldots0}^{11\ldots1} \eta_\gamma \sum_{\beta=00\ldots0}^{11\ldots1} a_{\beta\gamma} \sum_{\alpha=00\ldots0}^{i1\ldots1} k_{\beta\alpha}^0 L_\alpha = \sum_{\gamma=00\ldots0}^{11\ldots1} \eta_\gamma \sum_{\alpha=00\ldots0}^{11\ldots1} \left(\sum_{\beta=00\ldots0}^{11\ldots1} k_{\beta\alpha}^0 a_{\beta\gamma}\right) L_\alpha$$

which is equivalent to (17) since $k_{\beta\alpha}^0$ is the element in row α and column β of K_0^{-1}.

In order to obtain equation (18), we note that (25) and (26) imply $R = ((K_0')^{-1} \otimes I_N)L$ and $L = (K' \otimes I_N)M$, so that

$$R = ((K_0')^{-1}K' \otimes I_N)M$$

and we may write, similarly to (27) and (28),

$$R_\beta = \sum_{\alpha=00\ldots0}^{11\ldots1} k_{\beta\alpha}^* M_\alpha \tag{29}$$

where $k_{\beta\alpha}^*$ is the element in row β and column α of $(KK_0^{-1})'$. Upon substituting (29) into (16), we obtain

$$D = \sum_{\gamma=00\ldots0}^{11\ldots1} \eta_\gamma \sum_{\beta=00\ldots0}^{11\ldots1} a_{\beta\gamma} \sum_{\alpha=00\ldots0}^{11\ldots1} k_{\beta\alpha}^* M_\alpha = \sum_{\gamma=00\ldots0}^{11\ldots1} \eta_\gamma \sum_{\alpha=00\ldots0}^{11\ldots1} \left(\sum_{\beta=00\ldots0}^{11\ldots1} k_{\beta\alpha}^* a_{\beta\gamma} \right) M_\alpha$$

and (18) follows upon noting that $k_{\beta\alpha}^*$ is the element in row α and column β of KK_0^{-1}.

APPENDIX C. ESTIMATION OF THE VARIANCE COMPONENTS

This appendix may be viewed as a generalization of the first part of Amemiya [1] to the p-way classification; however, the method of proof is partly new and relies on Lemma 1. We first note that, since $\theta = K_0^{-1}\lambda$, unbiased estimators of the variance components are easy to obtain from unbiased estimators of the eigenvalues; however, since $\theta_{00\ldots0} = 0$ is known and since K_0^{-1} is upper triangular, we need only estimate the last $2^p - 1$ eigenvalues. For this reason we will always assume in this appendix that $\alpha \neq 00\ldots0$.

We will first present unbiased estimators $\bar{\lambda}_\alpha$ of the eigenvalues when the true disturbances are known. The $2^p - 1$ last equations of $\bar{\theta} = K_0^{-1}\bar{\lambda}$ then provide unbiased estimators of the variance components θ_α.

We will show that $\bar{\theta}_\alpha$ is consistent under the normality of u by proving that $V(\bar{\theta}_\alpha) \to 0$; the usual continuity argument does not apply since K_0^{-1} is a function of the numbers of observations n_i. We will then show that predicting the true disturbances by the analysis of covariance residuals (a procedure suggested by Amemiya) leads to estimated variance components $\hat{\theta}_\alpha$ with the same asymptotic distribution as the $\bar{\theta}_\alpha$. It follows that $\hat{\theta}_\alpha$ is a consistent, and asymptotically unbiased, estimator of θ_α.

In the omitted component case of Section 3, we let $\hat{\theta}_\gamma^* = 0$, and $\hat{\theta}_\alpha^* = \hat{\theta}_\alpha$ for $\alpha \neq \gamma$; misspecification does not affect the consistency of $\hat{\theta}_\alpha^*$ since λ is estimated from the correctly specified model.

We first show that when the true disturbances are known, unbiased estimators of the eigenvalues are given by $\bar{\lambda}_\alpha = \mu_\alpha^{-1} u' R_\alpha u$, where, from (14), $\mu_\alpha = \mathrm{tr}(R_\alpha) = \prod_{i=1}^{p} (n_i - 1)^{\alpha_i}$. This follows from Balestra [2, Lemma 1]:

$$E(\bar{\lambda}_\alpha) = \mu_\alpha^{-1} \mathrm{tr}(R_\alpha V) = \mu_\alpha^{-1} \mathrm{tr}\left(R_\alpha \sum_\beta \lambda_\beta R_\beta \right) = \lambda_\alpha \mu_\alpha^{-1} \mathrm{tr}(R_\alpha) = \lambda_\alpha.$$

Under the normality of u, we have, from Balestra [2, Lemma 6]:

$$V(u'R_\alpha u) = 2\,\mathrm{tr}(R_\alpha V R_\alpha V) = 2\,\mathrm{tr}(\lambda_\alpha^2 R_\alpha) = 2\lambda_\alpha^2 \mu_\alpha;$$

furthermore $u'R_\alpha u$ and $u'R_\beta u$ are independent for $\alpha \neq \beta$. Upon letting $\bar{\theta}_\alpha = \sum_\beta c_{\alpha\beta}\bar{\lambda}_\beta$ with $c_{\alpha\beta}$ the element in row α and column β of K_0^{-1}, we then have

$$V(\bar{\theta}_\alpha) = 2 \sum_\beta c_{\alpha\beta}^2 \mu_\beta^{-1} \lambda_\beta^2 = 2\lambda'\Delta\Delta_\alpha\lambda = 2\theta'K_0'\Delta\Delta_\alpha K_0\theta \tag{30}$$

where Δ and Δ_α are diagonal matrices with the μ_β^{-1} and the $c_{\alpha\beta}^2$ on the diagonal, respectively. It may be checked that

$$\Delta = \frac{1}{\prod\limits_{i=1}^{p} (n_i - 1)} \bigotimes_{i=1}^{p} \begin{pmatrix} n_i - 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{31}$$

$$\Delta_\alpha = \frac{1}{N^2} \bigotimes_{i=1}^{p} \begin{pmatrix} 0 & 0 \\ 0 & n_i^2 \end{pmatrix}^{\alpha_i} \tag{32}$$

so that the matrix in the quadratic form (30) may be written

$$K_0' \Delta \Delta_\alpha K_0 = \frac{1}{N^2 \prod\limits_{i=1}^{p} (n_i - 1)} \bigotimes_{i=1}^{p} \begin{pmatrix} n_i & 0 \\ 1 & 1 \end{pmatrix} \bigotimes_{i=1}^{p} \begin{pmatrix} n_i - 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\bigotimes_{i=1}^{p} \begin{pmatrix} 0 & 0 \\ 0 & n_i^2 \end{pmatrix}^{\alpha_i} \bigotimes_{i=1}^{p} \begin{pmatrix} n_i & 1 \\ 0 & 1 \end{pmatrix}$$

$$= \frac{1}{N^2 \prod\limits_{i=1}^{p} (n_i - 1)} \bigotimes_{i=1}^{p} \begin{pmatrix} 0 & 0 \\ 0 & n_i^2 \end{pmatrix}^{\alpha_i} \begin{pmatrix} n_i^2(n_i - 1) & n_i(n_i - 1) \\ n_i(n_i - 1) & n_i \end{pmatrix}^{1 - \alpha_i}$$

$$= \bigotimes_{i=1}^{p} \begin{pmatrix} 0 & 0 \\ 0 & (n_i - 1)^{-1} \end{pmatrix}^{\alpha_i} \begin{pmatrix} 1 & n_i^{-1} \\ n_i^{-1} & n_i^{-1}(n_i - 1)^{-1} \end{pmatrix}^{1 - \alpha_i} \stackrel{\text{def}}{=} P_\alpha. \tag{33}$$

It is easily seen from (33) that $\lim P_\alpha = O$, so that $V(\bar{\theta}_\alpha) \to 0$ in (30). Furthermore, it can be checked that $\lim \mu_\alpha P_\alpha = e_\alpha e_\alpha'$, where e_α is the $2^p \times 1$ unit vector with element α equal to unity; equation (30) then implies $\lim \mu_\alpha V(\bar{\theta}_\alpha) = 2\theta_\alpha^2$. This result, as well as equations (30) and (33), is consistent with Amemiya [1, pp. 5 and 6].

We now show the asymptotic equivalence when u is replaced by $\hat{u} = My = Mu$, where

$$M = I_N - \frac{1}{N} S_N - \left(I_N - \frac{1}{N} S_N\right) (X(X'QX)^{-1}X'Q)$$

with $Q \equiv R_{11\ldots 1} = \bigotimes_{i=1}^{p}(I_{n_i} - n_i^{-1}S_{n_i})$ (see Amemiya [1]). Since $R_\alpha S_N = S_N R_\alpha = O$, we have

$$M' R_\alpha M = R_\alpha - Q'X(X'QX)^{-1}X'R_\alpha X(X'QX)^{-1}X'Q.$$

Hence, upon letting

$$\hat{\theta}_\alpha = \sum_\beta c_{\alpha\beta} \hat{\lambda}_\beta = \sum_\beta c_{\alpha\beta} \mu_\beta^{-1} \hat{u}' R_\beta \hat{u}$$

we see that[4]

$$\sqrt{\mu_\alpha}(\hat{\theta}_\alpha - \theta_\alpha) = \sqrt{\mu_\alpha}(\bar{\theta}_\alpha - \theta_\alpha)$$

$$- \sqrt{\mu_\alpha} \sum_\beta c_{\alpha\beta} \mu_\beta^{-1} \frac{u'QX}{\sqrt{N}} \left(\frac{X'QX}{N}\right)^{-1} \left(\frac{X'R_\beta X}{N}\right) \left(\frac{X'QX}{N}\right)^{-1} \frac{X'Qu}{\sqrt{N}}. \tag{34}$$

The estimators $\hat{\theta}_\alpha$ and $\tilde{\theta}_\alpha$ will be asymptotically equivalent if the last terms in (34) converge in probability to zero. If $(X'QX)/N$ tends to a finite, positive definite matrix and if $(X'R_\beta X)/N$ is bounded, this will be true when $\sqrt{\mu_\alpha}\, c_{\alpha\beta}\mu_\beta^{-1}$ tends to 0 for all β, by the same argument as in Amemiya [1, equation (16)]. But $c_{\alpha\beta}\mu_\beta^{-1}$ is the element in row α and column β of $K_0^{-1}\Delta$, where Δ is defined in (31). Since, from (12) and (31),

$$K_0^{-1}\Delta = \frac{1}{N \prod\limits_{i=1}^{p}(n_i - 1)} \bigotimes_{i=1}^{p} \begin{pmatrix} 1 & -1 \\ 0 & n_i \end{pmatrix} \begin{pmatrix} n_i - 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{N} \bigotimes_{i=1}^{p} \begin{pmatrix} 1 & -(n_i - 1)^{-1} \\ 0 & n_i(n_i - 1)^{-1} \end{pmatrix},$$

we see that this is indeed the case.