

General Variance Covariance Structures in Two-Way Random Effects Models

Carlos De Porres, Jaya Krishnakumar

Department of Economics, University of Geneva, Geneva, Switzerland

Email: carlos.deporres@unige.ch, jaya.krishnakumar@unige.ch

Received January 11, 2013; revised February 22, 2013; accepted February 28, 2013

Copyright © 2013 Carlos De Porres, Jaya Krishnakumar. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

This paper examines general variance-covariance structures for the specific effects and the overall error term in a two-way random effects (RE) model. So far panel data literature has only considered these general structures in a one-way model and followed the approach of a Cholesky-type transformation to bring the model back to a “classical” one-way RE case. In this note, we first show that in a two-way setting it is impossible to find a Cholesky-type transformation when the error components have a general variance-covariance structure (which includes autocorrelation). Then we propose solutions for this general case using the spectral decomposition of the variance *components* and give a general transformation leading to a block-diagonal structure which can be easily handled. The results are obtained under some general conditions on the matrices involved which are satisfied by most commonly used structures. Thus our results provide a general framework for introducing new variance-covariance structures in a panel data model. We compare our results with [1] and [2] highlighting similarities and differences.

Keywords: Error Components; Matrix Decompositions; Panel Data; Spectral Decompositions

1. Introduction

Panel data models are often characterised by a three component error structure consisting of an individual (time invariant) specific effect, a time specific (individual invariant) effect and an overall idiosyncratic error term varying in both the individual and time dimensions. This leads to a corresponding three component variance covariance structure for the combined error term of the model. Due to the potentially large dimension of this variance covariance matrix, its inverse is usually calculated using matrix decomposition results. In a seminal paper [3] derived the spectral decomposition of the variance-covariance matrix of a two-way (and a one-way) random effects (error component-EC) model when all the components are assumed to be i.i.d. and independent two by two.

Several works have extended Nerlove’s result when non-i.i.d. structures are assumed for the error terms, namely, an MA(1) for the overall error term in [4], an AR(1) structure for the overall error term in [5], an ARMA structure for the overall error term in [6-8]. All these studies consider a one-way EC *i.e.* only individual

effect in addition to the overall error term, and apply a first-stage Cholesky-type transformation in order to get back to a “classical” one-way EC setting in which both the individual effect and the idiosyncratic error are i.i.d. after the transformation. Once we get back to the “classical” setting, we can apply Nerlove’s spectral decomposition for implementing the GLS procedure. Thus, all the studies so far have employed a combination of a first stage Cholesky and a second stage spectral decomposition.

In this note, we first show that the Cholesky-spectral combination approach adopted by all the above-mentioned articles is not possible in a two-way setting and there exists no transformation that will get us back to a “classical” EC structure. Then we propose a solution to the problem through a different first stage transformation based on a new spectral decomposition result which we present and prove. This new result is derived for two way EC models with general variance-covariance structures for all the error components. Our first stage transformation does not yield a classical two-way EC but a one-way EC with heteroscedastic errors and we show that the spectral decomposition and hence the determinant of the

variance covariance matrix of the transformed errors is easy to obtain, thus making GLS- and Maximum Likelihood-type procedures operational.

A recent article [1] provides a solution for the inverse and determinant of some general variance covariance structures in panel data models including a general one-way EC setting with a variance covariance matrix Ω of the form $\Gamma\Gamma' \otimes \Lambda\Lambda' + A_2 \otimes B_2'$ where Γ is a $T \times r_A$ matrix of full column rank, Λ is a $N \times r_B$ matrix of full column rank, A_2 and B_2 are positive definite matrices of order $T \times T$ and $N \times N$ respectively. They showed that a two-way EC of the form

$\Omega = aa' \otimes B + A \otimes bb' + \tau^2 A \otimes B$, where a is a T -vector, b is a N -vector, A is $T \times T$, and B is $N \times N$, is a special case of the above general one-way EC. However, the authors add that if one were to generalise the traditional three component structure, namely

$\Omega = A_1 \otimes B_1 + A_2 \otimes B_2 + A_3 \otimes B_3$, where all the first matrices of the Kronecker products are of order $T \times T$ and the second ones $N \times N$, then “the inverse is not easily computable in general”. The above general Kronecker structure is precisely what we consider in our note and we propose a general two-stage transformation for calculating its inverse and determinant under some conditions¹ Thus our result represents a natural follow up to the note by [1]. Our note *partially* validates their conjecture as our first stage spectral transformation yields a two component heteroscedastic structure and we do bring the model to a two component variance structure at an intermediate stage. However, we say only “partially” as we cannot *rewrite* our *original* variance matrix as a two component variance structure but apply a transformation to *convert* a general three component variance structure into a two component one.

Another more recent study [2] presents a general approach for a two-way EC under double autocorrelation in both the time effect and the idiosyncratic error. They propose a transformation based on a mixture of Cholesky and spectral decompositions as a first stage (what we call a “hybrid” transformation in this note), and a spectral decomposition as a second stage. In our note, we use our methodology to extend their approach to more general structures at the cross-sectional level. We derive the spectral decomposition as well as the determinant of the resulting variance-covariance matrix whereas [2] only derives the inverse of the transformed variance-covariance matrix.

The paper is organised as follows. Section 2 shows why Cholesky-type transformations do not work in the two-way EC model. In Section 3 we derive the new

“spectral-spectral” combination result for general structures of variance-covariance matrices. The new transformation is obtained under some conditions that are satisfied for most of the structures usually encountered in a panel data setting. Section 4 takes up some of these commonly used structures and describes how a solution can be found for these structures using our new decomposition result. Section 5 presents a new transformation along the lines of [2]. Finally, we conclude by pointing out some interesting aspects of our approach that may be worth investigating further.

2. On the Impossibility of Cholesky Decomposition in a Two-Way Error Component Context

Consider the following panel data model:

$$y_{it} = x'_{it}\beta + \omega_{it}, i = 1, \dots, N, t = 1, \dots, T \quad (1)$$

where

$$\omega_{it} = \mu_i + \lambda_t + \varepsilon_{it}$$

with μ_i denoting the individual specific random effect, λ_t denoting the time specific random effect, ε_{it} the overall idiosyncratic error term, and where β is a $k \times 1$ vector of coefficients including the intercept, i and t denote the individual and the time period respectively, x_{it} is a $k \times 1$ vector of observations on k strictly exogenous explanatory variables. ε_{it} is assumed to follow a stationary process $\forall t$ independent over $i = 1, \dots, N$, and parameterised by a vector θ . Writing

$$\mu = (\mu_1, \mu_2, \dots, \mu_N)', \lambda = (\lambda_1, \lambda_2, \dots, \lambda_T)',$$

$$\varepsilon = (\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1T}, \varepsilon_{21}, \dots, \varepsilon_{NT})'$$

let us assume that:

$$V(\lambda) = \sigma_\lambda^2 I_T, V(\mu) = \sigma_\mu^2 I_N, V(\varepsilon) = I_N \otimes \Omega_\varepsilon(\theta)$$

where $\Omega_\varepsilon(\theta)$ is any non-iid variance covariance structure and the disturbance components are independent two by two. Under these assumptions, the variance-covariance matrix of the linear model can be expressed as follows:

$$V(\omega) = \sigma_\mu^2 (I_N \otimes \mathbf{1}_T \mathbf{1}'_T) + \sigma_\lambda^2 (\mathbf{1}_N \mathbf{1}'_N \otimes I_T) + I_N \otimes \Omega_\varepsilon(\theta) \equiv V \quad (2)$$

In the case of a one-way EC model, the structure typically proposed in the literature for $\Omega_\varepsilon(\theta)$ is an autocorrelated structure and a Cholesky decomposition is used to get back to a “classical” one-way EC setting. In the following lemma, we establish that in a two-way EC model it is impossible to find a Cholesky decomposition that diagonalises $\Omega_\varepsilon(\theta)$ and gets us back to a “classical” setting. A solution to this problem is proposed in the

¹We show that these conditions are satisfied in almost all usually assumed structures and these conditions do not imply a rewriting of the three component variance matrix as a two component one. Thus we do go beyond a two component structure.

next section.

Lemma 1. Let $T = I_N \otimes T_\varepsilon$ be a transformation matrix, where T_ε is a $(T \times T)$ matrix such that $T_\varepsilon T_\varepsilon' = \Omega_\varepsilon(\theta)^{-1}$ with $\Omega_\varepsilon(\theta) \neq I_T$. Then, there is no matrix with the above-mentioned properties that allows us to get back to a classical two-way EC model.

Proof: See Appendix A.

3. Spectral Decompositions, an Alternative

This section shows that two-stage spectral decompositions can be applied to solve the problem of an analytical inverse of V for the general variance-covariance structure (2) presented in the previous section. Before deriving the main result in the form of a theorem, we illustrate our idea in the case of a two-way EC model with autocorrelation in the overall error term, a case that may be commonly encountered in practice (for which there is no explicit solution so far).

Assume that we are dealing with a panel data model with the following variance-covariance matrix (EC with MA structure for the idiosyncratic error):

$$V(\omega) = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\lambda^2(J_N \otimes I_T) + (I_N \otimes \Omega_{\varepsilon,b,T}) \quad (3)$$

where $\Omega_{\varepsilon,b,T} = \sigma_\varepsilon^2 \begin{bmatrix} 1+b^2 & b & 0 & \dots & 0 \\ b & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b \\ 0 & \dots & 0 & b & 1+b^2 \end{bmatrix}_{T \times T}$

As the inverse of the above variance-covariance matrix is rather difficult to compute, if we want to apply a GLS or Maximum Likelihood procedure we need to find a transformation which will yield a variance-covariance matrix with a tractable inverse. In Section 2, we proved that it was not possible to get to the classical structure through the Cholesky decomposition. Nor is it possible to rewrite the model as a one-way EC structure. Here we show that by using the spectral decomposition, we can provide an explicit transformation that allows us to get to a model with a variance-covariance matrix that is easy to handle.

First we note that [9] has given the orthogonal matrix $Q_{\varepsilon,b,T}$ such that $Q_{\varepsilon,b,T}' \Omega_{\varepsilon,b,T} Q_{\varepsilon,b,T} = \Lambda_{\varepsilon,b,T}$, where $\Lambda_{\varepsilon,b,T} = \text{diag}(\lambda_{\varepsilon,b,t})_{t=1,\dots,T}$, is the matrix of eigenvalues of $\Omega_{\varepsilon,b,T}$, with $\lambda_{\varepsilon,b,t} = 1 + b^2 - 2b \cos\left(\frac{t\pi}{T+1}\right)$. The t -th column of $Q_{\varepsilon,b,T}$ is given by:

$$\begin{aligned} & [Q_{\varepsilon,b,T}]'_{t=1,\dots,T} \\ & = \sqrt{\frac{2}{T+1}} \left[\sin\left(\frac{t\pi}{T+1}\right), \sin\left(\frac{2t\pi}{T+1}\right), \dots, \sin\left(\frac{Tt\pi}{T+1}\right) \right] \quad (4) \end{aligned}$$

Next we derive another result (Lemma 2 below) which gives the diagonalisation of a unit matrix and shows that the same orthogonal matrix also diagonalises an equicorrelation matrix as well. This can be particularly useful in panel data models with both individual and time specific effects. In fact, one of the main problems for getting the spectral decomposition of the full variance-covariance matrix in the presence of a time effect is the cross-sectional dependence induced by the latter. Indeed, in the presence of a time effect we lose the block-diagonal structure which is found in one-way models. This lemma is useful as it allows us to obtain a transformation leading to a block-diagonal structure.

Lemma 2. Let $\mathbf{1}_N$ be a $N \times 1$ vector of ones. Let $J_N = \mathbf{1}_N \mathbf{1}_N'$ and²

$$\Omega_{\mu,\rho,N} = \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \dots & \rho & 1 \end{bmatrix}_{N \times N}$$

Then the same matrix $Q_{i,N}$ diagonalises both J_N and $\Omega_{\mu,\rho,N}$ i.e.

$$\begin{aligned} Q_{i,N}' J_N Q_{i,N} &= \Lambda_{i,N} \\ Q_{i,N}' \Omega_{\mu,\rho,N} Q_{i,N} &= \Lambda_{\mu,\rho,N} \end{aligned}$$

where $Q_{i,N}$ represents the orthogonal matrix of eigenvectors of both J_N and $\Omega_{\mu,\rho,N}$ (its expression is given in the lemma's proof in the Appendix), and $\Lambda_{i,N}$ and $\Lambda_{\mu,\rho,N}$ represent the matrices of eigenvalues of J_N and $\Omega_{\mu,\rho,N}$ respectively.

Proof: See Appendix A.

Now combining [9]'s result and Lemma 2, we can give a transformation, $S_1 = Q_{i,N} \otimes Q_{\varepsilon,b,T}$ which when applied to (3) will lead to a block-diagonal structure whose spectral decomposition is easily obtained (see Theorem 1 and Special Case 1). Thus GLS and ML methods become much easier to operationalise in presence of complex variance covariance structures.

The method presented above fits in a more general approach that provides a way out for obtaining the inverse of any general variance-covariance structure for the specific effects as well as the idiosyncratic error in a two-way EC setting. We derive the solution under certain assumptions which may seem restrictive at first sight but which we show to be satisfied by many general structures frequently found in the literature. In fact we also show that this theorem enables us to introduce some new and possibly relevant structures in panel data models.

Let us now consider general processes for μ_i, λ_i and ε_{it} such that $V(\mu) = \Omega_\mu, V(\lambda) = \Omega_\lambda$, and $V(\varepsilon) = \Omega_\varepsilon$.

²Note that $\Omega_{\mu,\rho,N}$ given in this lemma represents the variance-covariance matrix of an equicorrelated error structure.

In other words we have individual (cross-sectional) dependence through Ω_μ , cross-sectional dependence and time dependence through Ω_λ and time dependence in Ω_ε .

Theorem 1. *Given a general variance covariance structure*

$$V(\omega) = \Omega_\mu \otimes J_T + J_N \otimes \Omega_\lambda + I_N \otimes \Omega_\varepsilon,$$

where $J_N = \iota_N \iota_N'$, if

$$\Omega_\mu J_N = J_N \Omega_\mu; \Omega_\lambda \Omega_\varepsilon = \Omega_\varepsilon \Omega_\lambda \tag{5}$$

then the following results hold³:

1) There exists a transformation matrix given by $Q_\omega = Q'_{\mu,t} \otimes Q'_{\lambda,\varepsilon}$ such that:

$$V^r(Q_\omega \omega) = \tilde{D}^r \begin{bmatrix} \tilde{D}_\mu \otimes \left(\frac{\tilde{d}\tilde{d}'}{\delta} \right) + I_{N-1} \otimes \left(I_T - \frac{\tilde{d}\tilde{d}'}{\delta} \right) & \mathbf{0} \\ \mathbf{0} & (\tilde{\delta}d_{\mu,N} + 1)^r \frac{\tilde{d}\tilde{d}'}{\tilde{\delta}} + I_T - \frac{\tilde{d}\tilde{d}'}{\tilde{\delta}} \end{bmatrix} \tilde{D}^r \tag{7}$$

where $r \in \mathbb{R}$, $\tilde{D}_\mu = \text{diag}(\delta d_{\mu,i} + 1)_{i=1, \dots, N-1}$, $d_{\mu,i}$ is the i -th element of the diagonal matrix D_μ , $\delta = \tilde{d}\tilde{d}'$, $\tilde{d} = D_\varepsilon^{-\frac{1}{2}} d$,

$$\begin{aligned} Q'_{\mu,t} \Omega_\mu Q_{\mu,t} &= D_\mu \\ Q'_{\mu,t} J_N Q_{\mu,t} &= D_t \\ Q'_{\lambda,\varepsilon} \Omega_\lambda Q_{\lambda,\varepsilon} &= D_\lambda \\ Q'_{\lambda,\varepsilon} \Omega_\varepsilon Q_{\lambda,\varepsilon} &= D_\varepsilon \end{aligned} \tag{6}$$

$$V(Q_\omega \omega) = D_\mu \otimes dd' + \tilde{D}$$

where $\tilde{D} = \begin{bmatrix} I_{N-1} \otimes D_\varepsilon & \mathbf{0} \\ \mathbf{0} & ND_\lambda + D_\varepsilon \end{bmatrix}$, $Q_{\mu,t}$ and $Q_{\lambda,\varepsilon}$ are orthogonal matrices, $d = Q'_{\lambda,\varepsilon} \iota_T$, and $D_\mu, D_t, D_\lambda, D_\varepsilon, \tilde{D}$ are diagonal matrices.

2) The spectral decomposition of the transformed variance-covariance matrix is given by:

$$\tilde{\delta} = \tilde{d}\tilde{d}' \text{ and } \tilde{d} = (ND_\lambda + D_\varepsilon)^{-\frac{1}{2}} d.$$

3) The determinant of the transformed variance-covariance matrix can be written as:

$$|V(Q_\omega \omega)| = \prod_{j=1}^T d_{\varepsilon,j}^{N-1} \prod_{k=1}^T (Nd_{\lambda,k} + d_{\varepsilon,k}) \prod_{i=1}^{N-1} (\delta d_{\mu,i} + 1) (\tilde{\delta}d_{\mu,N} + 1) \tag{8}$$

Proof: See Appendix B.

Remarks:

1) By assuming variance-covariance matrices of a general form for the three components in the random effects model, denoted as $\Omega_\mu, \Omega_\lambda$ and Ω_ε , we go further than any other article that has dealt with finding appropriate transformations for non i.i.d. error structures in a panel data framework. All explicit transformations that exist in the literature are given for $\Omega_\lambda = \mathbf{0}, \Omega_\mu = \sigma_\mu^2 I_N$ or $\text{diag}(\sigma_{\mu,i}^2)_{i=1, \dots, N}$ and Ω_ε issued from an AR(1) and/or MA(1) process. All these features can be seen as special cases of the result stated in Theorem 1 (see Section 4).

2) Our results are derived under the double commutativity constraint (5). Though this condition looks rather restrictive, it turns out to be quite general in the panel data framework. In fact, not only all the structures studied so far in the panel econometric literature satisfy this condition but also there are other possibly interesting structures which have not yet been considered that can be included in this setting. One such structure is equicorre-

lation which can be particularly relevant for individual effects in certain empirical contexts. It can even be modified to reflect block dependence among individuals (correlation within clusters).

3) We show different ways of taking cross-sectional dependence into account, *i.e.*, through Ω_μ and/or Ω_λ . The presence of the time effect is already a source of cross-sectional dependence in our model and it can be further generalised to have a fuller variance-covariance structure through Ω_λ . This feature can be linked to the more recent strand of literature that deals with cross-sectional dependence through factor analysis.

4) Although our result is general, it may present some operational difficulties. Our transformation matrix requires the knowledge of eigenvectors and eigenvalues of the different matrices involved and it may be cumbersome to actually determine these for some structures. Unlike the MA case that we saw earlier, in the case of an AR structure for λ or ε , there are no general analytical expressions available for the eigenvectors of the resulting variance covariance matrix for any dimension T . These expressions crucially depend on the size of T . Plus there is no recursive way of finding the roots of characteristic polynomials of size $T + 1$ given those of size T

³We will come back to the commutativity condition in a remark following the theorem (see Remark 2).

and one has to calculate them separately for each T . These are practical obstacles to be overcome before implementing our transformation. In spite of this, we believe that our transformation is highly useful as it definitely reduces the size of the problem in all circumstances *i.e.* instead of finding the eigenvalues/eigenvectors of a matrix of size NT (*i.e.* of the full variance-covariance matrix of the model), one only needs to find them for a matrix of size T (*i.e.* of the AR structure over just the time dimension). In the worst case scenario, one can compute these eigenvalues and eigenvectors numerically in the first stage.

4. Some Special Cases

In this section we will examine some special cases that are commonly used for one way models and show how our result enables to extend them to the two-way case and obtain the spectral decomposition of the variance covariance matrix of the model in a rather straightforward manner. In addition, we also introduce some new structures that have not been considered so far in this context.

1) Independent μ_i 's, independent λ_i 's, and MA(1) ε_{it} 's with the same parameter b for all i 's. In other words $\Omega_\mu = \sigma_\mu^2 I_N, \Omega_\lambda = \sigma_\lambda^2 I_T$ and $\Omega_\varepsilon = \Omega_{\varepsilon,b,T}$ of an MA(1) process. This is the case that we already considered in the beginning of Section 3.

One can easily verify that the commutativity condition is satisfied in this case *i.e.* $\Omega_\mu J_N = J_N \Omega_\mu$ and $\Omega_\lambda \Omega_\varepsilon = \Omega_\varepsilon \Omega_\lambda$ ⁴.

Then, the variance-covariance of the whole model writes:

$$V_1 = \sigma_\mu^2 (I_N \otimes J_T) + \sigma_\lambda^2 (J_N \otimes I_T) + (I_N \otimes \Omega_{\varepsilon,b,T}) \quad (9)$$

Using Theorem 1, the transformation to be applied to get a block diagonal structure is given by $S_1 = Q'_{i,N} \otimes Q'_{\varepsilon,b,T}$, with $Q_{i,N}$ given in the Appendix and the matrix of eigenvectors of $\Omega_{\varepsilon,b,T}$ previously defined. Note that this transformation does not depend on any nuisance parameter, so it can be applied as a first stage transformation without any prior estimation. The variance-covariance matrix of the transformed error, \tilde{V}_1 then becomes:

$$\tilde{V}_1 = \sigma_\mu^2 (I_N \otimes d_1 d_1') + \sigma_\lambda^2 (\Lambda_{i,N} \otimes I_T) + I_N \otimes \Lambda_{\varepsilon,b,T} \quad (10)$$

with $d_1 = Q'_{\varepsilon,b,T} t_T$.

One can easily see that \tilde{V}_1 can be written in the form $D_\mu \otimes dd' + \tilde{D}$ denoting $\sigma_\mu^2 I_N$ as D_μ and the diagonal matrix $\sigma_\lambda^2 (\Lambda_{i,N} \otimes I_T) + I_N \otimes \Lambda_{\varepsilon,b,T}$ as \tilde{D} . Then the spectral decomposition of \tilde{V}_1 can be easily computed using the result in point 2 of Theorem 1.

2) Equicorrelated μ_i 's, independent λ_i 's, and MA(1)

ε_{it} 's with the same parameter b for all i 's. In other words, $\Omega_\mu = \Omega_{\mu,\rho}, \Omega_\lambda = \sigma_\lambda^2 I_T$ and $\Omega_\varepsilon = \Omega_{\varepsilon,b,T}$.

$$V_2 = \Omega_{\mu,\rho} \otimes J_T + \sigma_\lambda^2 (J_N \otimes I_T) + I_N \otimes \Omega_{\varepsilon,b,T} \quad (11)$$

Here one can use the same transformation as in the previous point (due to Lemma 2), *i.e.*, $S_2 = Q'_{i,N} \otimes Q'_{\varepsilon,b,T}$. The corresponding transformed variance-covariance matrix is:

$$\tilde{V}_2 = \Lambda_{\mu,\rho,N} \otimes d_1 d_1' + \sigma_\lambda^2 (\Lambda_{i,N} \otimes I_T) + I_N \otimes \Lambda_{\varepsilon,b,T} \quad (12)$$

As in case 1, the spectral decomposition of the transformed variance-covariance matrix can be explicitly computed using the result in point 2 of Theorem 1 denoting $\Lambda_{\mu,\rho,N}$ as D_μ and the diagonal matrix $\sigma_\lambda^2 (\Lambda_{i,N} \otimes I_T) + I_N \otimes \Lambda_{\varepsilon,b,T}$ as \tilde{D} .

Note that the transformation for the above 2 cases does not exist in the literature. It is the first time, to our knowledge, that one can deal with cross-sectional dependence and autocorrelated errors over time in the same transformation in panel data models.

3) Equicorrelated μ_i 's, MA(1) λ_i 's, and independent ε_{it} 's. In other words, $\Omega_\mu = \Omega_{\mu,\rho}, \Omega_\lambda = \Omega_{\lambda,b,T}$ and $\Omega_\varepsilon = \sigma_\varepsilon^2 I_T$.

This case is symmetric to the previous one and hence, the transformation remains the same, $S_3 = Q'_{i,N} \otimes Q'_{\lambda,b,T}$, the original variance-covariance being:

$$V_3 = \Omega_{\mu,\rho} \otimes J_T + J_N \otimes \Omega_{\lambda,b,T} + \sigma_\varepsilon^2 (I_N \otimes I_T) \quad (13)$$

The variance-covariance matrix of the transformed model is:

$$\tilde{V}_3 = \Lambda_{\mu,\rho,N} \otimes d_1 d_1' + \Lambda_{i,N} \otimes \Lambda_{\lambda,b,T} + I_N \otimes I_T \quad (14)$$

Once again noting that the first term of \tilde{V}_3 is of the form $D_\mu \otimes dd'$ and the second plus the third term is a diagonal matrix of the form \tilde{D} , one gets the spectral decomposition of \tilde{V}_3 from Theorem 1.

4) Equicorrelated μ_i 's, equicorrelated λ_i 's, and equicorrelated ε_{it} 's. That is $\Omega_\mu = \Omega_{\mu,\rho,N}, \Omega_\lambda = \Omega_{\lambda,\rho,T}$ and $\Omega_\varepsilon = \Omega_{\varepsilon,\rho,T}$.

Here we assume both cross-sectional and time equicorrelation. The transformation matrix is given by $S_4 = Q'_{i,N} \otimes Q'_{i,T}$. The original variance-covariance matrix is:

$$V_4 = \Omega_{\mu,\rho,N} \otimes J_S + J_N \otimes \Omega_{\lambda,\rho,T} + I_N \otimes \Omega_{\varepsilon,\rho,T} \quad (15)$$

and the transformed variance-covariance matrix is:

$$\tilde{V}_4 = \Lambda_{\mu,\rho,N} \otimes d_1 d_1' + \Lambda_{i,N} \otimes \Lambda_{\lambda,\rho,T} + I_N \otimes \Lambda_{\varepsilon,\rho,T} \quad (16)$$

The spectral decomposition of \tilde{V}_4 follows from point 2, Theorem 1, as in the above cases.

5) Equicorrelated or i.i.d μ_i 's, i.i.d λ_i 's and AR(1) ε_{it} 's.

Referring to our Remark 4 of Section 3, the eigenvectors of an AR(1) variance covariance matrix are different for different T 's and cumbersome to calculate. Although

⁴All the special cases discussed below satisfy this condition and hence we will not mention it each time.

we have them for $T = 5$ for example, we do not present them here as the expressions of the eigenvectors are indeed long. One can also calculate them numerically. Once the first stage decomposition is obtained (perhaps numerically for the third AR component), the procedure in Theorem 1 can be implemented as in the above cases.

5. Extension to “Hybrid” Transformations

In spite of the fact that the commutativity constraint is verified in many panel data models of practical relevance, one could argue that it is stringent and may exclude some potentially important situations. For instance, assume that the researcher expects to have an autocorrelated process not only in the idiosyncratic error but also in the time effect. In general, the variance-covariance matrices of the two autocorrelated processes do not commute. [2] proposes a 3-stage transformation that circumvents the double autocorrelation problem. They successively apply the Cholesky and the spectral transformations to obtain a simpler structure for which they are able to give the inverse. Thus their first two stages are equivalent to our first stage in the sense that after two transformations they provide a way to calculate the inverse of the resulting variance-covariance matrix. However, they do not provide the spectral decomposition of the Cholesky transformed variance-covariance matrix in the first stage as it cannot be explicitly derived.

In this section we examine these more general cases and give appropriate transformations based on our previous results. These transformations are characterised by

$$V^r(H\omega) = \mathbf{D}_{\mu,H}^{\frac{r}{2}} \begin{bmatrix} \tilde{\mathbf{D}}_{\mu,H} \otimes \left(\frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\delta_{\varepsilon,c,\lambda}} \right) + \mathbf{I}_{N-1} \otimes \left(\mathbf{I}_T - \frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\delta_{\varepsilon,c,\lambda}} \right) & \mathbf{0} \\ \mathbf{0} & \left(\tilde{\delta}_{\varepsilon,c,\lambda} \mathbf{d}_{\mu,N,H} + 1 \right)^r \frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\tilde{\delta}_{\varepsilon,c,\lambda}} + \left(\mathbf{I}_T - \frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\tilde{\delta}_{\varepsilon,c,\lambda}} \right) \mathbf{D}_{\mu,H}^{\frac{r}{2}} \end{bmatrix} \quad (19)$$

where $r \in \mathbb{R}$, $\tilde{\mathbf{D}}_H = \text{diag}(\delta_{\varepsilon,c,\lambda} d_{\mu,i,H} + 1)_{i=1,\dots,N-1}$, d_{μ,i,H_1} is the i -th element of the diagonal matrix \mathbf{D}_μ , $\delta_{\varepsilon,c,\lambda} = \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}_{\varepsilon,c,\lambda}$, $\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} = \mathbf{D}_{\varepsilon,c,\lambda}^{-\frac{1}{2}} \mathbf{d}_{\varepsilon,c,\lambda}$, $\tilde{\delta}_{\varepsilon,c,\lambda} = \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}_{\varepsilon,c,\lambda}$

$$|V(H\omega)| = \prod_{j=1}^T d_{\mu,j,H}^{N-1} \prod_{k=1}^T (N + d_{\mu,i,H}) \prod_{i=1}^{N-1} (\delta_{\varepsilon,c,\lambda} d_{\mu,i,H} + 1) (\tilde{\delta}_{\varepsilon,c,\lambda} d_{\mu,N,H} + 1) \quad (20)$$

Proof: See Appendix C.

Remarks:

1) Note that Theorem 2 considers a slightly more general structure than [2] by introducing cross-sectional de-

a mixture of Cholesky and spectral decompositions (what we call “hybrid” transformations) in the first stage and a spectral decomposition in the second stage. The following theorem provides a new transformation in this general setting that allows us to obtain the spectral decomposition of the resulting variance-covariance matrix.

Theorem 2. Given a general variance-covariance structure

$$V(\omega) = \mathbf{\Omega}_\mu \otimes \mathbf{J}_T + \mathbf{J}_N \otimes \mathbf{\Omega}_\lambda + \mathbf{I}_N \otimes \mathbf{\Omega}_\varepsilon,$$

if

$$\mathbf{\Omega}_\mu \mathbf{J}_N = \mathbf{J}_N \mathbf{\Omega}_\mu \quad (17)$$

then the following results hold:

1) There exists a transformation matrix of the form $\mathbf{H} = \mathbf{Q}'_{\mu,t} \otimes (\mathbf{Q}'_{\varepsilon,c,\lambda} \mathbf{C}'_\lambda)$ such that:

$$\begin{aligned} \mathbf{Q}'_{i,\mu} \mathbf{\Omega}_\mu \mathbf{Q}_{i,\mu} &= \mathbf{D}_\mu \\ \mathbf{Q}'_{i,\mu} \mathbf{J}_N \mathbf{Q}_{i,\mu} &= \mathbf{D}_i \\ \mathbf{C}'_\lambda \mathbf{\Omega}_\lambda \mathbf{C}_\lambda &= \mathbf{I}_T \\ \mathbf{Q}'_{\varepsilon,c,\lambda} (\mathbf{C}'_\lambda \mathbf{\Omega}_\varepsilon \mathbf{C}_\lambda) \mathbf{Q}_{\varepsilon,c,\lambda} &= \mathbf{D}_{\varepsilon,c,\lambda} \\ V(\mathbf{H}\omega) &= \mathbf{D}_\mu \otimes \mathbf{d}_{c,\lambda} \mathbf{d}'_{c,\lambda} + \mathbf{D}_H \end{aligned} \quad (18)$$

where $\mathbf{D}_H = \begin{bmatrix} \mathbf{I}_{N-1} \otimes \mathbf{D}_{\varepsilon,c,\lambda} & \mathbf{0} \\ \mathbf{0} & N\mathbf{I}_T + \mathbf{D}_{\varepsilon,c,\lambda} \end{bmatrix}$, $\mathbf{Q}_{\mu,t}$ and $\mathbf{Q}_{\varepsilon,c,\lambda}$ are orthogonal matrices, $\mathbf{d}_{c,\lambda} = \mathbf{Q}'_{\varepsilon,c,\lambda} \mathbf{C}'_\lambda \mathbf{1}_T$, and $\mathbf{D}_\mu, \mathbf{D}_i, \mathbf{D}_{\varepsilon,c,\lambda}$ and \mathbf{D}_H are diagonal matrices.

2) The spectral decomposition of the transformed variance-covariance matrix is given by:

and $\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} = (N\mathbf{I}_T + \mathbf{D}_{\varepsilon,c,\lambda})^{-\frac{1}{2}} \mathbf{d}_{\varepsilon,c,\lambda}$.

3) The determinant of the transformed variance-covariance matrix is given by:

pendence through $\mathbf{\Omega}_\mu$ in addition to that induced by the time effect.

2) The transformations given in point 1 of the theorem are called “hybrid” since they use both spectral, $\mathbf{Q}_{\mu,t}$

and $\mathbf{Q}_{\varepsilon,c,\lambda}$, and Cholesky, \mathbf{C}_λ , decompositions in the first stage in order to arrive at a tractable variance-covariance matrix.

3) These transformations are more general than those of Theorem 1 in the sense that they can handle more complex structures in $\mathbf{\Omega}_\lambda$ and $\mathbf{\Omega}_\varepsilon$, in particular autocorrelation of the AR type. However, the problem of finding such transformations becomes even more difficult than in Section 3. Indeed, even if we have an idea of the Cholesky decomposition of the variance-covariance matrix of a particular stochastic process, say $\mathbf{\Omega}_\lambda$, such knowledge is not, in general, informative for the spectral decomposition of the variance-covariance matrix of another stochastic process, $\mathbf{\Omega}_\varepsilon$, nor for a linear transformation of the latter, $\mathbf{C}'_\lambda \mathbf{\Omega}_\varepsilon \mathbf{C}_\lambda$. Yet, we require all these decompositions to implement the procedure.

4) The results of this section are only theoretical as we cannot provide any explicit transformation, especially in the first stage, as mentioned above. However, our earlier Theorem 1 provides an explicit solution under the commutativity constraint.

5) We still have the commutativity constraint between $\mathbf{\Omega}_\mu$ and \mathbf{J}_N in our case that extends [2]'s model to cover general variance covariance structures at the cross-sectional level⁵. We conjecture that in the two-way EC this is the maximum generality that one can afford in order to get back to a heteroscedastic one-way case after the first transformation.

6. Concluding Remarks

In this paper, we examine general variance-covariance structures for the specific effects and the overall error term in a two-way error component model. We show the limitations of Cholesky-type transformations and propose a different approach, based on spectral decomposition, for dealing with time *and* cross-sectional dependence. Our transformation can be applied to any general variance-covariance setting, under the commutativity constraint, and we show how this transformation works in many interesting special cases. We also connect our result to [1] and their conjecture for a two-way EC which seems to be verified albeit only after an initial *transformation*.

As our transformation is based on eigenvalues and eigenvectors of the variance-covariance matrix of the combined disturbance term, we believe that it is strongly linked to the more recent approach of taking cross-sectional dependence into account by means of factor mod-

els. We conjecture that our approach is equivalent to the factor approach under some assumptions and we hope to investigate the link between these two approaches in the future.

Finally, we show how the result derived in [2] for the double autocorrelation structure can be extended to cover more general structures at the cross-sectional level using our method. We provide the spectral decomposition as well as the determinant of the variance-covariance matrix of the transformed model.

REFERENCES

- [1] J. R. Magnus and C. Muris, "Specification of Variance Matrices for Panel Data Models," *Econometric Theory*, Vol. 26, No. 1, 2010, pp. 301-310. [doi:10.1017/S0266466609090756](https://doi.org/10.1017/S0266466609090756)
- [2] J. M. B. Brou, E. Kouassi and K. O. Kymn, "Double Autocorrelation in Two Way Error Component Models," *Open Journal of Statistics*, Vol. 1, No. 3, 2011, pp. 185-198. [doi:10.4236/ojs.2011.13022](https://doi.org/10.4236/ojs.2011.13022)
- [3] M. Nerlove, "A Note on Error Components Models," *Econometrica*, Vol. 39, No. 2, 1971, pp. 383-396. [doi:10.2307/1913351](https://doi.org/10.2307/1913351)
- [4] P. Balestra, "A Note on the Exact Transformation Associated with the First-Order Moving Average Process," *Journal of Econometrics*, Vol. 14, No. 3, 1980, pp. 381-394. [doi:10.1016/0304-4076\(80\)90034-2](https://doi.org/10.1016/0304-4076(80)90034-2)
- [5] B. H. Baltagi and Q. Li, "A Transformation That Will Circumvent the Problem of Autocorrelation in an Error Component Model," *Journal of Econometrics*, Vol. 48, No. 3, 1991, pp. 385-393. [doi:10.1016/0304-4076\(91\)90070-T](https://doi.org/10.1016/0304-4076(91)90070-T)
- [6] V. Zinde-Walsh, "Some Exact Formulae for Autoregressive Moving Average Processes," *Econometric Theory*, Vol. 4, No. 3, 1988, pp. 384-402. [doi:10.1017/S0266466600013360](https://doi.org/10.1017/S0266466600013360)
- [7] J. W. Galbraith and V. Zinde-Walsh, "The GLS Transformation Matrix and a Semi-Recursive Estimator for the Linear Regression Model with ARMA Errors," *Econometric Theory*, Vol. 8, No. 1, 1992, pp. 95-111. [doi:10.1017/S0266466600010756](https://doi.org/10.1017/S0266466600010756)
- [8] J. W. Galbraith and V. Zinde-Walsh, "Transforming the Error-Components Model for Estimation with General ARMA Disturbances," *Journal of Econometrics*, Vol. 66, No. 1-2, 1995, pp. 349-355. [doi:10.1016/0304-4076\(94\)01621-6](https://doi.org/10.1016/0304-4076(94)01621-6)
- [9] M. H. Pesaran, "Exact Maximum Likelihood Estimation of a Regression Equation with a First Order Moving Average Errors," *The Review of Economic Studies*, Vol. 40, No. 124, 1973, pp. 529-538. [doi:10.2307/2296586](https://doi.org/10.2307/2296586)

⁵In their model $\mathbf{\Omega}_\mu = \mathbf{I}_N$ and hence this condition is automatically satisfied.

Appendix A

Proof of Lemma 1. *It is straightforward that,*

$$V(\mathbf{T}\omega) = \sigma_\mu^2 (\mathbf{I}_N \otimes \mathbf{T}_\varepsilon \mathbf{t}_T \mathbf{t}'_T \mathbf{T}'_\varepsilon) + \sigma_\lambda^2 (\mathbf{t}_N \mathbf{t}'_N \otimes \mathbf{T}_\varepsilon \mathbf{T}'_\varepsilon) + \mathbf{I}_N \otimes \mathbf{T}_\varepsilon \boldsymbol{\Omega}_\varepsilon(\theta) \mathbf{T}'_\varepsilon \quad (21)$$

In order to get back to a classical EC structure we would need two conditions:

- 1) $\mathbf{T}_\varepsilon \boldsymbol{\Omega}_\varepsilon(\theta) \mathbf{T}'_\varepsilon = \mathbf{I}_T$
- 2) $\mathbf{T}_\varepsilon \mathbf{T}'_\varepsilon = \mathbf{I}_T$

However, condition 1 implies $\mathbf{T}_\varepsilon \mathbf{T}'_\varepsilon = \boldsymbol{\Omega}_\varepsilon^{-1}(\theta)$. This in turn implies $\boldsymbol{\Omega}_\varepsilon^{-1}(\theta) = \mathbf{I}_T$. Hence the only matrix $\boldsymbol{\Omega}_\varepsilon(\theta)$ that satisfies both conditions in a two-EC is the

$$\mathbf{P}_{i,N} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 & 1 \\ -\frac{1}{N-1} & 1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & -\frac{1}{N-2} & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & -\frac{1}{N-3} & \ddots & 0 & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 & 1 \\ -\frac{1}{N-1} & -\frac{1}{N-2} & -\frac{1}{N-3} & \dots & -\frac{1}{2} & 1 \end{bmatrix},$$

$$\mathbf{O}_{i,N} = \begin{bmatrix} \sqrt{\frac{N-1}{N}} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{\frac{N-2}{N-1}} & 0 & \dots & 0 \\ \vdots & \ddots & \sqrt{\frac{N-3}{N-2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & \frac{1}{\sqrt{N}} \end{bmatrix},$$

- 2) Diagonalization of $\boldsymbol{\Omega}_{\mu,\rho,N}$

Same reasoning as in the previous point. $\boldsymbol{\Omega}_{\mu,\rho,N}$ is symmetric, therefore there exists an orthogonal matrix such that $\mathbf{Q}'_{i,N} \boldsymbol{\Omega}_{\mu,\rho,N} \mathbf{Q}_{i,N} = \boldsymbol{\Lambda}_{\mu,\rho,N}$ where $\boldsymbol{\Lambda}_{\mu,\rho,N}$, the matrix of eigenvalues, is

$$\boldsymbol{\Lambda}_{\mu,\rho,N} = \begin{bmatrix} (1-\rho) \mathbf{I}_{(N-1) \times (N-1)} & \mathbf{0}_{(N-1) \times 1} \\ \mathbf{0}_{1 \times (N-1)} & 1 + (N-1)\rho \end{bmatrix}$$

Direct calculation shows that $\mathbf{Q}_{i,N}$ of point 1 satisfies the above condition. ■

Appendix B

Proof of Theorem 1. *The variance-covariance of the transformed error term can be written as:*

identity matrix meaning i.i.d. errors. In this case there is no need for implementing any transformation since we are already in the classical two-way-EC case⁶. ■

Proof of Lemma 2.

- 1) Diagonalization of \mathbf{J}_N

Since \mathbf{J}_N is symmetric, there exists an orthogonal matrix, $\mathbf{Q}_{i,N}$, such that $\mathbf{Q}'_{i,N} \boldsymbol{\Lambda}_{i,N} \mathbf{Q}_{i,N} = \mathbf{J}_N$, where $\boldsymbol{\Lambda}_{i,N}$, the matrix of eigenvalues, is

$$\boldsymbol{\Lambda}_{i,N} = \begin{bmatrix} \mathbf{0}_{(N-1) \times (N-1)} & \mathbf{0}_{(N-1) \times 1} \\ \mathbf{0}_{1 \times (N-1)} & N \end{bmatrix}$$

It can be easily verified that the solution for $\mathbf{Q}_{i,N}$ is given by $\mathbf{Q}_{i,N} = \mathbf{P}_{i,N} \mathbf{O}_{i,N}$ with

$$V(\mathbf{Q}\omega) = \underbrace{\mathbf{Q}'_{\mu,t} \boldsymbol{\Omega}_\mu \mathbf{Q}_{\mu,t}}_{\mathbf{D}_\mu} \otimes \underbrace{\mathbf{Q}'_{\lambda,\varepsilon} \mathbf{t}_T \mathbf{t}'_T \mathbf{Q}_{\lambda,\varepsilon}}_{\mathbf{d}} + \underbrace{\mathbf{Q}'_{\mu,t} \mathbf{J}_N \mathbf{Q}_{\mu,t}}_{\mathbf{D}_t} \otimes \underbrace{\mathbf{Q}'_{\lambda,\varepsilon} \boldsymbol{\Omega}_\lambda \mathbf{Q}_{\lambda,\varepsilon}}_{\mathbf{D}_\lambda} + \underbrace{\mathbf{Q}'_{\mu,t} \mathbf{Q}_{\mu,t}}_{\mathbf{I}_N} \otimes \underbrace{\mathbf{Q}'_{\lambda,\varepsilon} \boldsymbol{\Omega}_\varepsilon \mathbf{Q}_{\lambda,\varepsilon}}_{\mathbf{D}_\varepsilon} \quad (22)$$

Proof of point 1.

From Lemma 2 we have $\mathbf{D}_t = \begin{bmatrix} \mathbf{0}_{(N-1) \times (N-1)} & \mathbf{0}_{(N-1) \times 1} \\ \mathbf{0}_{1 \times (N-1)} & qN \end{bmatrix}$

⁶One can show that the only possibility that a Cholesky type transformation will allow us to get back to a classical EC will be when λ_t and ε_{it} follow the same stochastic process. As this hypothesis is very stringent, one needs to find new transformations in the general case.

where $q \in \mathbb{R}$. Hence,

$$\begin{aligned} & \mathbf{D}_i \otimes \mathbf{D}_\lambda + \mathbf{I}_N \otimes \mathbf{D}_\varepsilon \\ &= \begin{bmatrix} \mathbf{I}_{N-1} \otimes \mathbf{D}_\varepsilon & \mathbf{0} \\ \mathbf{0} & qN\mathbf{D}_\lambda + \mathbf{D}_\varepsilon \end{bmatrix} \equiv \tilde{\mathbf{D}}. \end{aligned}$$

This completes the proof of the first point.

Proof of point 2.

$$V(\mathbf{Q}_\omega \omega) = \tilde{\mathbf{D}}^{\frac{1}{2}} \underbrace{\left(\tilde{\mathbf{D}}^{-\frac{1}{2}} (\mathbf{D}_\mu \otimes d d') \tilde{\mathbf{D}}^{-\frac{1}{2}} + \mathbf{I}_{NT} \right)}_{\mathbf{V}_D} \tilde{\mathbf{D}}^{\frac{1}{2}} \quad (23)$$

where

$$\begin{aligned} \mathbf{V}_D &= \begin{bmatrix} \mathbf{D}_{\mu, [N-1]} \otimes \underbrace{\mathbf{D}_\varepsilon^{\frac{1}{2}} d d' \mathbf{D}_\varepsilon^{\frac{1}{2}}}_d & \mathbf{0} \\ \mathbf{0} & d_{\mu, N} \underbrace{(qN\mathbf{D}_\lambda + \mathbf{D}_\varepsilon)^{\frac{1}{2}} d d' (qN\mathbf{D}_\lambda + \mathbf{D}_\varepsilon)^{\frac{1}{2}}}_d \end{bmatrix} + \mathbf{I}_{NT} \\ &= \begin{bmatrix} \mathbf{D}_{\mu, [N-1]} \otimes \tilde{d} \tilde{d}' & \mathbf{0} \\ \mathbf{0} & d_{\mu, N} \tilde{d} \tilde{d}' \end{bmatrix} + \mathbf{I}_{NT} = \begin{bmatrix} \delta \mathbf{D}_{\mu, [N-1]} \otimes \frac{\tilde{d} \tilde{d}'}{\delta} + \mathbf{I}_{N-1} \otimes \mathbf{I}_T & \mathbf{0} \\ \mathbf{0} & d_{\mu, N} \delta \frac{\tilde{d} \tilde{d}'}{\delta} + \mathbf{I}_T \end{bmatrix} \end{aligned} \quad (24)$$

with $\mathbf{D}_{\mu, [N-1]} = \text{diag}_{i=1, \dots, N-1} (d_{\mu, i})$. \mathbf{V}_D is a block diagonal matrix. Hence we can compute the spectral decomposition for each block and this will give rise to the spectral decomposition of the whole matrix. For the first $N-1$ blocks, we have:

$$[\mathbf{V}_D]_i = \delta d_{\mu, i} \frac{\tilde{d} \tilde{d}'}{\delta} + \mathbf{I}_T = (\delta d_{\mu, i} + 1) \frac{\tilde{d} \tilde{d}'}{\delta} + \mathbf{I}_T - \frac{\tilde{d} \tilde{d}'}{\delta}, \quad (25)$$

$i = 1, \dots, N-1$.

and similarly for

$$\mathbf{V}_D^r = \begin{bmatrix} \text{diag}(\delta d_{\mu, i} + 1)_{i=1, \dots, N-1}^r \otimes \frac{\tilde{d} \tilde{d}'}{\delta} + \mathbf{I}_{N-1} \otimes \left(\mathbf{I}_T - \frac{\tilde{d} \tilde{d}'}{\delta} \right) & \mathbf{0} \\ \mathbf{0} & (\delta d_{\mu, N} + 1)^r \frac{\tilde{d} \tilde{d}'}{\delta} + \mathbf{I}_T - \frac{\tilde{d} \tilde{d}'}{\delta} \end{bmatrix} \quad (26)$$

Therefore we have

$$V(\mathbf{Q}_\omega \omega)^r = \tilde{\mathbf{D}}^{\frac{r}{2}} \mathbf{V}_D^r \tilde{\mathbf{D}}^{\frac{r}{2}} \quad (27)$$

Proof of point 3.

$$|V(\mathbf{Q}_\omega \omega)| = \left| \tilde{\mathbf{D}}^{\frac{1}{2}} \right| |V_D| \left| \tilde{\mathbf{D}}^{\frac{1}{2}} \right| = |\tilde{\mathbf{D}}| |V_D| \quad (28)$$

where

$$|\tilde{\mathbf{D}}| = |\mathbf{D}_\varepsilon|^{N-1} |N\mathbf{D}_\lambda + \mathbf{D}_\varepsilon| = \prod_{j=1}^T d_{\varepsilon, j}^{N-1} \prod_{k=1}^T (Nd_{\lambda, k} + d_{\varepsilon, k}) \quad (29)$$

For $|V_D|$, we know that the determinant of a block-diagonal matrix is equal to the product of the determinant of each block. Since the eigenvalues for each block are given in point 2 (spectral decomposition), the determinant of each block is equal to the product of the eigen-

values. Hence, $[\mathbf{V}_D]_N = (\delta d_{\mu, N} + 1) \frac{\tilde{d} \tilde{d}'}{\delta} + \mathbf{I}_T - \frac{\tilde{d} \tilde{d}'}{\delta}$. As $\frac{\tilde{d} \tilde{d}'}{\delta}$ and $\mathbf{I}_T - \frac{\tilde{d} \tilde{d}'}{\delta}$ are idempotent, orthogonal to each other and their sum is equal to the identity matrix, the last two expressions give the spectral decompositions for each block with $(\delta d_{\mu, i} + 1)$ and 1 being the eigenvalues for blocks $i = 1, \dots, N-1$, and $(\delta d_{\mu, N} + 1)$ and 1 the eigenvalues for block N . Hence,

values. Hence,

$$|V_D| = (\delta d_{\mu, N} + 1) \prod_{i=1}^{N-1} (\delta d_{\mu, i} + 1) \quad (30)$$

This completes the proof of Theorem 1.

Appendix C

Proof of Theorem 2. The variance-covariance of the transformed error term is:

$$\begin{aligned} V(\mathbf{H}\omega) &= \underbrace{\mathbf{Q}'_{\mu, i} \mathbf{Q}_\mu \mathbf{Q}_{\mu, i}}_{\mathbf{D}_\mu} \otimes \underbrace{\mathbf{Q}'_{\varepsilon, c, \lambda} \mathbf{C}'_\lambda \mathbf{I}_T \mathbf{I}'_T \mathbf{C}_\lambda \mathbf{Q}_{\varepsilon, c, \lambda}}_{\mathbf{d}_{c, \lambda}} \\ &+ \underbrace{\mathbf{Q}'_{\mu, i} \mathbf{J}_N \mathbf{Q}_{\mu, i}}_{\mathbf{D}_i} \otimes \underbrace{\mathbf{Q}'_{\varepsilon, c, \lambda} \mathbf{C}'_\lambda \mathbf{Q}_\lambda \mathbf{C}_\lambda \mathbf{Q}_{\varepsilon, c, \lambda}}_{\mathbf{I}_T} \\ &+ \underbrace{\mathbf{Q}'_{\mu, i} \mathbf{Q}_{\mu, i}}_{\mathbf{I}_N} \otimes \underbrace{\mathbf{Q}'_{\varepsilon, c, \lambda} \mathbf{C}'_\lambda \mathbf{Q}_\varepsilon \mathbf{C}_\lambda \mathbf{Q}_{\varepsilon, c, \lambda}}_{\mathbf{D}_{\varepsilon, c, \lambda}} \end{aligned} \quad (31)$$

Proof of point 1.

As in point 1 of Theorem 1 we have

$$\begin{aligned} & \mathbf{D}_t \otimes \mathbf{I}_T + \mathbf{I}_N \otimes \mathbf{D}_{\varepsilon,c,\lambda} \\ &= \begin{bmatrix} \mathbf{I}_{N-1} \otimes \mathbf{D}_{\varepsilon,c,\lambda} & \mathbf{0} \\ \mathbf{0} & N\mathbf{I}_T + \mathbf{D}_{\varepsilon,c,\lambda} \end{bmatrix} \equiv \mathbf{D}_H. \end{aligned}$$

This completes the proof of point 1.

Proof of point 2.

We proceed in the same way as in point 2 of Theorem 1.

$$V(\mathbf{H}\omega) = \mathbf{D}_H^{\frac{1}{2}} \underbrace{\left(\mathbf{D}_H^{-\frac{1}{2}} (\mathbf{D}_\mu \otimes \mathbf{d}_{\varepsilon,c,\lambda} \mathbf{d}'_{\varepsilon,c,\lambda}) \mathbf{D}_H^{-\frac{1}{2}} + \mathbf{I}_{NT} \right)}_{\mathbf{V}_H} \mathbf{D}_H^{\frac{1}{2}} \quad (32)$$

$$\begin{aligned} \mathbf{V}_H &= \begin{bmatrix} \mathbf{D}_{\mu,[N-1]} \otimes \underbrace{\mathbf{D}_{\varepsilon,c,\lambda}^{-\frac{1}{2}} \mathbf{d}_{\varepsilon,c,\lambda} \mathbf{d}'_{\varepsilon,c,\lambda}}_{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda}} & \mathbf{0} \\ \mathbf{0} & d_{\mu,N,H} \underbrace{(\mathbf{N}\mathbf{I}_T + \mathbf{D}_{\varepsilon,c,\lambda})^{-\frac{1}{2}} \mathbf{d}_{\varepsilon,c,\lambda} \mathbf{d}'_{\varepsilon,c,\lambda}}_{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda}} \end{bmatrix} + \mathbf{I}_{NT} \\ &= \begin{bmatrix} \delta_{\varepsilon,c,\lambda} \mathbf{D}_{\mu,[N-1]} \otimes \frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\delta_{\varepsilon,c,\lambda}} + \mathbf{I}_{N-1} \otimes \mathbf{I}_T & \mathbf{0} \\ \mathbf{0} & d_{\mu,N,H} \tilde{\delta}_{\varepsilon,c,\lambda} \frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\tilde{\delta}_{\varepsilon,c,\lambda}} + \mathbf{I}_T \end{bmatrix} \end{aligned} \quad (33)$$

Applying the same argument as in the previous theorem, we get:

$$\mathbf{V}_H^r = \begin{bmatrix} \tilde{\mathbf{D}}_{\mu,H}^r \otimes \left(\frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\delta_{\varepsilon,c,\lambda}} \right) + \mathbf{I}_{N-1} \otimes \left(\mathbf{I}_T - \frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\delta_{\varepsilon,c,\lambda}} \right) & \mathbf{0} \\ \mathbf{0} & (\tilde{\delta}_{\varepsilon,c,\lambda} d_{\mu,N,H} + 1)^r \frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\tilde{\delta}_{\varepsilon,c,\lambda}} + \left(\mathbf{I}_T - \frac{\tilde{\mathbf{d}}_{\varepsilon,c,\lambda} \tilde{\mathbf{d}}'_{\varepsilon,c,\lambda}}{\tilde{\delta}_{\varepsilon,c,\lambda}} \right) \end{bmatrix} \quad (34)$$

Hence,

$$V(\mathbf{H}\omega) = \mathbf{D}_H^{\frac{r}{2}} \mathbf{V}_H^r \mathbf{D}_H^{\frac{r}{2}} \quad (35)$$

and

Proof of point 3.

$$|V(\mathbf{H}\omega)| = |\mathbf{D}_H|^r |\mathbf{V}_H| \quad (36)$$

$$|\mathbf{V}_H| = \prod_{j=1}^{N-1} (\delta_{\varepsilon,c,\lambda} d_{\mu,j,H} + 1) (\tilde{\delta}_{\varepsilon,c,\lambda} d_{\mu,N,H} + 1)$$

where

This completes the proof of Theorem 2.