Estimating Sparse Variance-covariance using Shrinkage Methods: Applications to Error Components Models

F. Chan [a] and **R. Chariag** [b]

[a] *School of Accounting, Economics and Finance, Curtin University, GPO BOX U1987, Perth, Western Australia, 6845*

[b] *Department of Economics, Central European University, Quellenstraße 51, A-1100 Wien, Austria, Vienna Commercial Court, FN 502313 x*

Email: chariag_ramzi@phd.ceu.edu

Abstract: This paper aims to investigate if shrinkage estimation of sparse variance-covariance matrix can be helpful in producing more robust feasible generalised least squares estimator (FGLS) in the regression context. Consider

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{X}$ and $\mathbf{y}$ are the $N \times K$ data matrix containing the data from K explanatory variables and $N \times 1$ vector containing the data of the dependent variable, respectively, while $\boldsymbol{\beta}$ is the $K \times 1$ vector of unknown parameter vector. The FGLS estimator is defined as

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{y} \tag{2}$$

where $\hat{\boldsymbol{\Omega}}$ denotes the estimated variance-covariance matrix of $\boldsymbol{\varepsilon}$. The computation of $\hat{\boldsymbol{\Omega}}$ in practice is usually based on the context of the model and/or assumptions underlying the idiosyncratic shocks, $\boldsymbol{\varepsilon}$. The validity of these assumptions is often difficult to verify in practice. The objective of this paper is to investigate to what extend can shrinkage estimation of $\boldsymbol{\Omega}$ be helpful in reducing the assumptions required to estimate $\boldsymbol{\Omega}$.

This paper will first consider the benchmark model in the form of a three-dimensional panel data model:

$$y_{ijt} = \mathbf{x}'_{ijt}\boldsymbol{\beta} + \underbrace{\alpha_i + \gamma_j + \lambda_t + u_{ijt}}_{\varepsilon_{ijt}}, \qquad i = 1, \ldots, N_1,\ j = 1, \ldots, N_2,\ t = 1, \ldots, T, \tag{3}$$

where α_i, γ_j, λ_t and u_{ijt} are pair-wise independent. Under the assumptions that each of these random variables are independently and identically distributed, the variance-covariance matrix of ε_{ijt} is

$$\boldsymbol{\Omega} = \sigma_\alpha^2\mathbf{S}_\alpha\mathbf{S}'_\alpha + \sigma_\gamma^2\mathbf{S}_\gamma\mathbf{S}'_\gamma + \sigma_\lambda^2\mathbf{S}_\lambda\mathbf{S}'_\lambda + \sigma_u^2\mathbf{I}_N \tag{4}$$

where $N = N_1N_2T$ with σ_α^2, σ_γ^2, σ_λ^2 and σ_u^2 denote the variance of α_i, γ_j, λ_t and u_{ijt}, respectively and $\mathbf{S}_\alpha$, $\mathbf{S}_\gamma$ and $\mathbf{S}_\lambda$ denote $\mathbf{I}_{N_1} \otimes \mathbf{i}_{N_2} \otimes \mathbf{i}_T$, $\mathbf{i}_{N_1} \otimes \mathbf{I}_{N_2} \otimes \mathbf{i}_T$ and $\mathbf{i}_{N_1} \otimes \mathbf{i}_{N_2} \otimes \mathbf{I}_T$, respectively.

The variance-covariance matrix $\boldsymbol{\Omega}$ is not a diagonal matrix in this case even though the variance-covariance matrix of each of the error components is a diagonal matrix. It is, however, a sparse matrix. The structure of the matrix depends on the specification of the error components as well as their variance-covariance structure, which is not known in practice. This paper will investigate the effectiveness of shrinkage estimation of $\boldsymbol{\Omega}$ in the context of FGLS across different assumptions on the error components.

Keywords: *M-estimator, initial values, optimisation, artificial neural network, machine learning*