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

Ramzi Chariag, Felix Chan

Outline

Motivation and Research Question

- In multi-dimensional panels, the number of possible specifications in the error components grows exponentially as the number of dimensions increases.
- Different specifications and independence assumptions lead to different sparsity structure of their variance-covariance matrix.
- It is possible to examine possible misspecification in the error components by leveraging specific sparsity structures.

Research Question

1. **Theory:** How can we leverage different ways to estimate covariance matrices in the context of a multidimensional panel?
2. **Useful Applications**:
    - The same collection of stocks in different markets, observed over multiple periods.
    - The same students taking many different subjects (that are also the same over time), observed over multiple periods.

Literature

There is a somewhat large literature on covariance matrix estimation. I make use of, and contribute to two specifically:

- **Maximum Likelihood Estimators**: (Meng 1991), (Wei 1990),(kuk 1997),(Pourahmadi, 2000)
- **Shrinkage Estimators**: (Fan 2008), Fan(2018), (Bien and Tibshirani 2011), (Ledoit and Wolf 2003),(Ledoit and Wolf 2023), (Mattera 2023)

(Bickel and Levina, 2008): Regularise a matrix by imposing constraints on its eigenvalues. The restriction is essentially that the maximum and minimum eigenvalues of the covariance matrix are close.

Set up

The number of possible combinations of fixed effects specifications increases grows exponentially, but we consider a simple example of a misspecification. Consider the following model:

$$\begin{aligned} y_{ijt} =& x'_{ijt}\beta + u_{ijt} \\ u_{ijt} =& \alpha_i + \gamma_j + \lambda_t + \varepsilon_{ijt} \end{aligned}$$

$$i = 1, \ldots, N_1, \, j = 1, \ldots, N_2, \;\; t = 1, \ldots, T,$$

We start off thinking about these vectors as fixed, but later we will instead fix their distribution. We then try to estimate the covariance matrix of that distribution.

β Estimators

- We will estimate the model using plain OLS, then compare it to GLS estimations. We use two estimators for Ω, which generate two $FGLS$ estimators.

$$
\begin{aligned}
\beta_{OLS} =& (X'X)^{-1}X'y \\
\beta_{GLS} =& (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y
\end{aligned}
$$

Different Specifications.

$$
\begin{aligned}
w_{ijt} =& \alpha_i + \gamma_j + \lambda_t \\
w_{ijt} =& \alpha_i + \delta_{jt} \\
w_{ijt} =& \omega_{it} + \delta_{jt} \\
w_{ijt} =& \pi_{ij} + \delta_{jt} + \omega_{it}.
\end{aligned}
$$

Corresponding Ω matrices (If homoschedastic, constant diagonal):

$$
\begin{aligned}
& S_\alpha \Omega_\alpha S_\alpha' + S_\gamma \Omega_\gamma S_\gamma' + S_\lambda \Omega_\lambda S_\lambda' + \sigma_u^2 I_N \\
& S_\alpha \Omega_\alpha S_\alpha' + S_\delta \Omega_\delta S_\delta' + \sigma_u^2 I_N \\
& S_\omega \Omega_\omega S_\omega' + S_\delta \Omega_\delta S_\delta' + \sigma_u^2 I_N \\
& S_\pi \Omega_\pi S_\pi + S_\delta \Omega_\delta S_\delta' + S_\omega \Omega_\omega S_\omega' + \sigma_u^2 I_N
\end{aligned}
$$

Core Estimation Idea

If the observations are sorted such that the innermost dimension of ordering is i, then j then t, and we are in the case of $w_{ijt} = \omega_{it} + \delta_{jt}$ the resulting covariance matrix of $\hat{u}_{ijt}$, be it Ω, would look something like:

$$\begin{bmatrix} \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} & 0 & 0 & 0 \\ 0 & \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} & 0 & 0 \\ 0 & 0 & \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} & 0 \\ 0 & 0 & 0 & \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \end{bmatrix}$$

Ω Estimators

We estimate Ω_c using an outer product of the residuals by cluster.

- Each block constitutes a cluster c.
- Let Ω_c be the covariance block of cluster c, and ε_c the vector of residuals associated with the same cluster.

$$\hat{\Omega}_c = \varepsilon_c \varepsilon_c'$$

Ω Estimators cont.

Outer products are symmetric and positive semi-definite by construction.

Problem: Outer products are singular.

Solution: $\hat{\Omega}_d = \frac{1}{C}\sum_{c=1}^{C}\Omega_c$, Where d is dimension, c is cluster, and C is the number of clusters. Let D be the number of observations within dimension d. As long as $C \geq D$, the matrix is non-singular.

FGLS1

Recall: $\hat{\Omega}_c = \varepsilon_c \varepsilon_c'$. All ε vectors in this case are assumed to be *iid*. We just take an outer product of all of them and average directly.

FGLS2

We first extract our vectors:

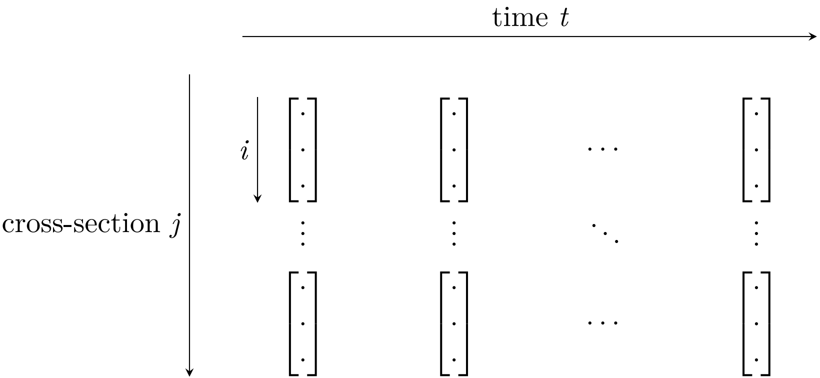


Figure: Panel of the same vector observed over time t, and across another dimension j.

FGLS2 cont.

Then we follow a two step procedure, depending on our modelling choice.
Example: The vector of i random effects is either redrawn across t, and observed with noise over j. In the first case, the estimation procedure becomes:

1. Average element by element over j, to obtain $\bar{\delta}_i$ from δ_{ij}
2. Average the outer product of our mean vectors.

Summary

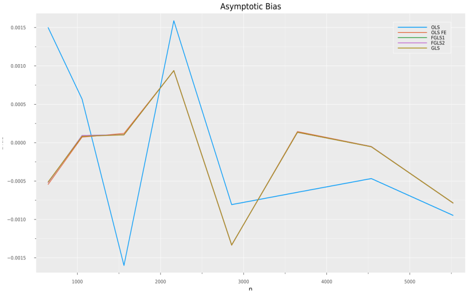
We will be looking at the performance of the following estimators:

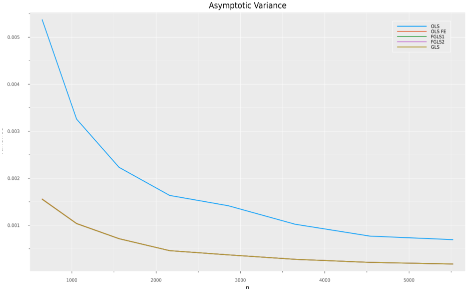
1. **Plain OLS**
2. **OLS FE**
3. **FGLS 1 & 2**: $\hat{\Omega}_d = \frac{1}{C}\sum_{c=1}^{C}\Omega_c$, with different cluster structures.
4. **Oracle GLS**: Benchmark for FGLS2.

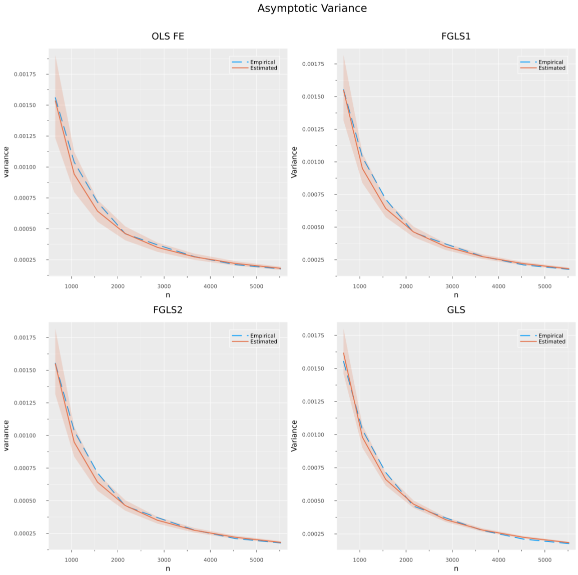
In the following possibilities:

- The vectors can either be fixed, or redrawn.
- Covariance between the vectors is either homoschedastic, or has a full covariance block.

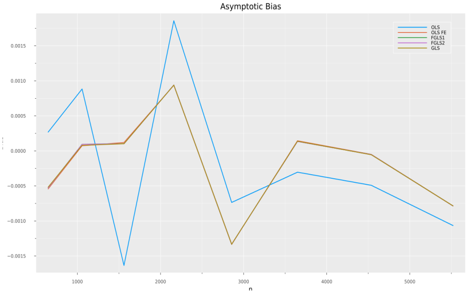
Simulation Results FE: Homoschedastic

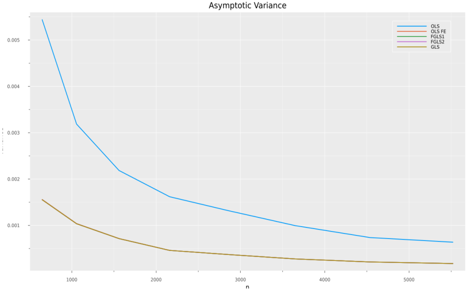


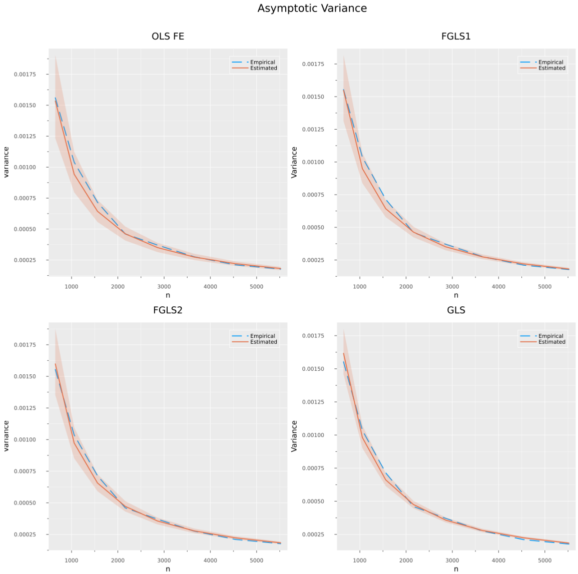




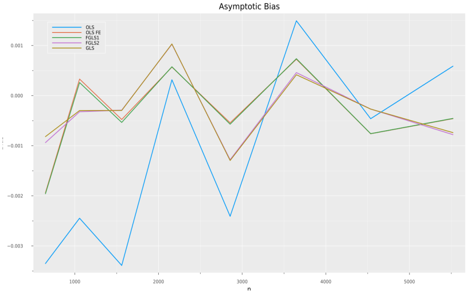
Simulation Results FE: Full i Covariance Block

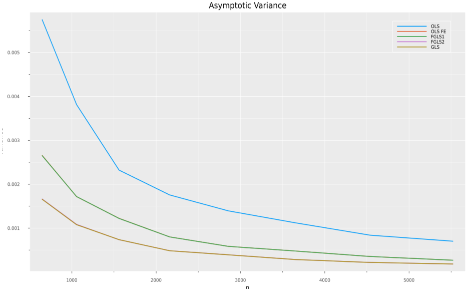


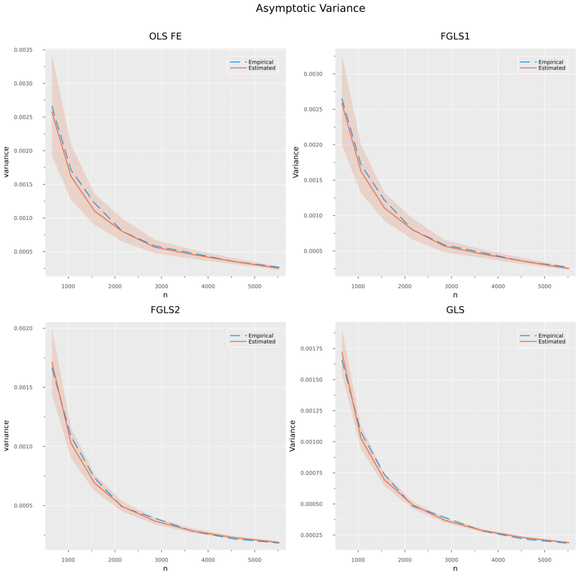




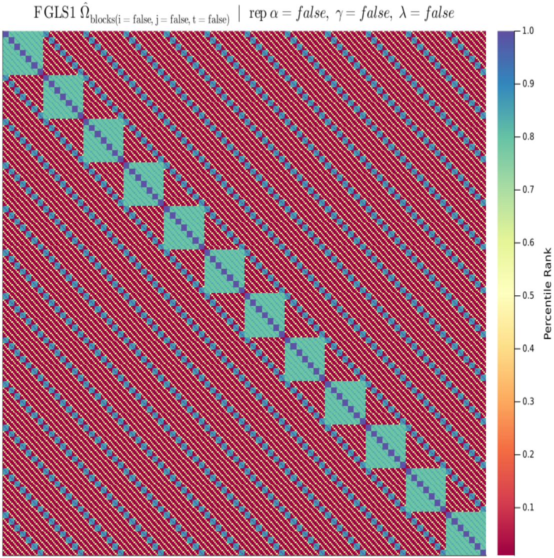
Simulation Results RE i: Full i Covariance Block

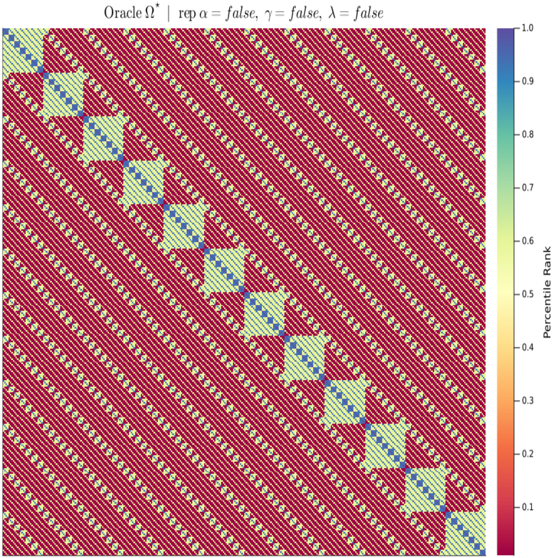




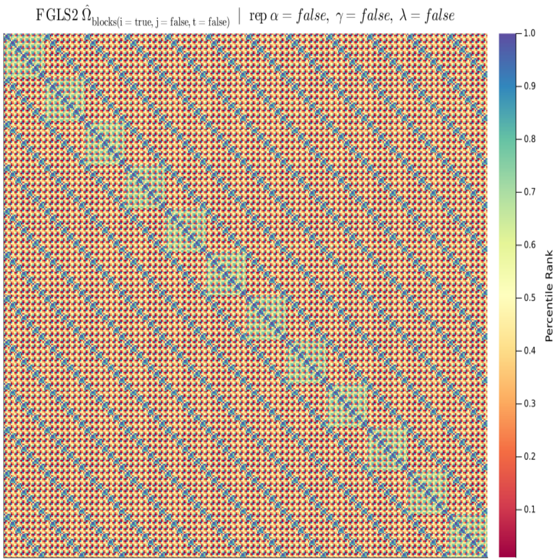


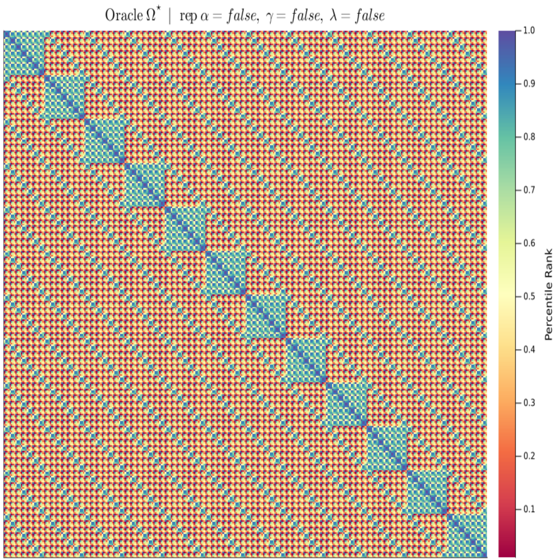
Example Covariance Structures: Homoschedastic Threeway FE





Example Covariance Structures: i Block Threeway FE





Example Covariance Structures: i RE