Why Panel Data Models?

We can follow several units over time, therefore we can get:

- Difference between units
- Development over time
- Efficiency gains
  - More observations
  - Seemingly Unrelated Regression (SUR)
  - Group-wise heteroskedasticity
  - Time-wise autoregression

What makes the difference from Time series data or Cross-sectional data? Two approaches in literature>

1. Analysis error term ε
2. Analysis of coefficients β
   - Fixed effects - βs are different for each individual
   - Random effects - the differences are “random”

Panel Balanced
Unbalanced

Short run adjustment $\Delta y_i = \beta^T \Delta X_i + \Delta\varepsilon_i$

Long run adjustment $y_i = \beta^T X_i + \varepsilon_i$

Pooling the data

Suppose **N** individuals(groups) - i, **T** time points - t, **K** regressors

General form of the equation:

$$Y_{it} = \beta_1 X_{it,1} + \beta_2 X_{it,2} + \beta_3 X_{it,3} + \dots + \beta_K X_{it,K} + \varepsilon_{it}$$

Stacking observations:

$$Y_{11} = \beta_1 X_{11,1} + \beta_2 X_{11,2} + \beta_3 X_{11,3} + \ldots + \beta_K X_{11,K} + \varepsilon_{11}$$
$$Y_{12} = \beta_1 X_{12,1} + \beta_2 X_{12,2} + \beta_3 X_{12,3} + \ldots + \beta_K X_{12,K} + \varepsilon_{12}$$
$$Y_{13} = \beta_1 X_{13,1} + \beta_2 X_{13,2} + \beta_3 X_{13,3} + \ldots + \beta_K X_{13,K} + \varepsilon_{13}$$

….

$$Y_{1T} = \beta_1 X_{1T,1} + \beta_2 X_{1T,2} + \beta_3 X_{1T,3} + \ldots + \beta_K X_{1T,K} + \varepsilon_{1T}$$
$$Y_{21} = \beta_1 X_{21,1} + \beta_2 X_{21,2} + \beta_3 X_{21,3} + \ldots + \beta_K X_{21,K} + \varepsilon_{21}$$
$$Y_{22} = \beta_1 X_{22,1} + \beta_2 X_{22,2} + \beta_3 X_{22,3} + \ldots + \beta_K X_{22,K} + \varepsilon_{22}$$

….

$$Y_{2T} = \beta_1 X_{2T,1} + \beta_2 X_{2T,2} + \beta_3 X_{2T,3} + \ldots + \beta_K X_{2T,K} + \varepsilon_{2T}$$

…

…

…

$$Y_{N1} = \beta_1 X_{N1,1} + \beta_2 X_{N1,2} + \beta_3 X_{N1,3} + \ldots + \beta_K X_{N1,K} + \varepsilon_{N1}$$

…

$$Y_{NT} = \beta_1 X_{NT,1} + \beta_2 X_{NT,2} + \beta_3 X_{NT,3} + \ldots + \beta_K X_{NT,K} + \varepsilon_{NT}$$

Matrix form:

$$\boldsymbol{y} = \boldsymbol{X}\ \boldsymbol{\beta} + \varepsilon$$

where

$$\boldsymbol{y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ \vdots \\ Y_{1T} \\ Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2T} \\ \vdots \\ Y_{N1} \\ \vdots \\ Y_{NT} \end{bmatrix} \quad \boldsymbol{X} = \begin{bmatrix} X_{11,1} & X_{11,2} & \cdots & X_{11,K} \\ X_{12,1} & X_{12,2} & \cdots & X_{12,K} \\ X_{13,1} & X_{13,2} & \cdots & X_{13,K} \\ \vdots & \vdots & \cdots & \vdots \\ X_{1T,1} & X_{1T,2} & \cdots & X_{1T,K} \\ X_{21,1} & X_{21,2} & \cdots & X_{21,K} \\ X_{22,1} & X_{22,2} & \cdots & X_{22,K} \\ \vdots & \vdots & \cdots & \vdots \\ X_{2T,1} & X_{2T,2} & \cdots & X_{2T,K} \\ \vdots & \vdots & \cdots & \vdots \\ X_{N1,1} & X_{N1,2} & \cdots & X_{N1,K} \\ \vdots & \vdots & \cdots & \vdots \\ X_{NT,1} & X_{NT,2} & \cdots & X_{NT,K} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \vdots \\ \varepsilon_{1T} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2T} \\ \vdots \\ \varepsilon_{N1} \\ \vdots \\ \varepsilon_{NT} \end{bmatrix}$$

NT x 1	NT x K	K x 1	NT x 1

Dimensions

Variance-covariance matrix E ($\varepsilon\,\varepsilon^T$)

Dimension is **<u>NT x NT</u>**

$$E(\varepsilon\varepsilon^T)=\begin{bmatrix} E(\varepsilon_{11}\varepsilon_{11}) & E(\varepsilon_{11}\varepsilon_{12}) & \dots & E(\varepsilon_{11}\varepsilon_{1T}) & E(\varepsilon_{11}\varepsilon_{21}) & \dots & E(\varepsilon_{11}\varepsilon_{NT}) \\ E(\varepsilon_{12}\varepsilon_{11}) & E(\varepsilon_{12}\varepsilon_{12}) & \dots & E(\varepsilon_{12}\varepsilon_{1T}) & E(\varepsilon_{12}\varepsilon_{21}) & \dots & E(\varepsilon_{12}\varepsilon_{NT}) \\ E(\varepsilon_{13}\varepsilon_{11}) & E(\varepsilon_{13}\varepsilon_{12}) & \dots & E(\varepsilon_{13}\varepsilon_{1T}) & E(\varepsilon_{13}\varepsilon_{21}) & \dots & E(\varepsilon_{13}\varepsilon_{NT}) \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ E(\varepsilon_{1T}\varepsilon_{11}) & E(\varepsilon_{1T}\varepsilon_{12}) & \dots & E(\varepsilon_{1T}\varepsilon_{1T}) & E(\varepsilon_{1T}\varepsilon_{21}) & \dots & E(\varepsilon_{1T}\varepsilon_{NT}) \\ E(\varepsilon_{21}\varepsilon_{11}) & E(\varepsilon_{21}\varepsilon_{12}) & \dots & E(\varepsilon_{21}\varepsilon_{1T}) & E(\varepsilon_{21}\varepsilon_{21}) & \dots & E(\varepsilon_{21}\varepsilon_{NT}) \\ E(\varepsilon_{22}\varepsilon_{11}) & E(\varepsilon_{22}\varepsilon_{12}) & \dots & E(\varepsilon_{22}\varepsilon_{1T}) & E(\varepsilon_{22}\varepsilon_{21}) & \dots & E(\varepsilon_{22}\varepsilon_{NT}) \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ E(\varepsilon_{2T}\varepsilon_{11}) & E(\varepsilon_{2T}\varepsilon_{12}) & \dots & E(\varepsilon_{2T}\varepsilon_{1T}) & E(\varepsilon_{2T}\varepsilon_{21}) & \dots & E(\varepsilon_{2T}\varepsilon_{NT}) \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ E(\varepsilon_{N1}\varepsilon_{11}) & E(\varepsilon_{N1}\varepsilon_{12}) & \dots & E(\varepsilon_{N1}\varepsilon_{1T}) & E(\varepsilon_{N1}\varepsilon_{21}) & \dots & E(\varepsilon_{N1}\varepsilon_{NT}) \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ E(\varepsilon_{NT}\varepsilon_{11}) & E(\varepsilon_{NT}\varepsilon_{12}) & \dots & E(\varepsilon_{NT}\varepsilon_{1T}) & E(\varepsilon_{NT}\varepsilon_{21}) & \dots & E(\varepsilon_{NT}\varepsilon_{NT}) \end{bmatrix}$$

$$
E(\varepsilon\varepsilon^{T}) = \begin{bmatrix}
E(\varepsilon_{11}^2) & E(\varepsilon_{11}\varepsilon_{12}) & \dots & E(\varepsilon_{11}\varepsilon_{1T}) & E(\varepsilon_{11}\varepsilon_{21}) & \dots & E(\varepsilon_{11}\varepsilon_{NT}) \\
E(\varepsilon_{12}\varepsilon_{11}) & E(\varepsilon_{12}^2) & \dots & E(\varepsilon_{12}\varepsilon_{1T}) & E(\varepsilon_{12}\varepsilon_{21}) & \dots & E(\varepsilon_{12}\varepsilon_{NT}) \\
E(\varepsilon_{13}\varepsilon_{11}) & E(\varepsilon_{13}\varepsilon_{12}) & \dots & E(\varepsilon_{13}\varepsilon_{1T}) & E(\varepsilon_{13}\varepsilon_{21}) & \dots & E(\varepsilon_{13}\varepsilon_{NT}) \\
\vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
E(\varepsilon_{1T}\varepsilon_{11}) & E(\varepsilon_{1T}\varepsilon_{12}) & \dots & E(\varepsilon_{1T}^2) & E(\varepsilon_{1T}\varepsilon_{21}) & \dots & E(\varepsilon_{1T}\varepsilon_{NT}) \\
E(\varepsilon_{21}\varepsilon_{11}) & E(\varepsilon_{21}\varepsilon_{12}) & \dots & E(\varepsilon_{21}\varepsilon_{1T}) & E(\varepsilon_{21}^2) & \dots & E(\varepsilon_{21}\varepsilon_{NT}) \\
E(\varepsilon_{22}\varepsilon_{11}) & E(\varepsilon_{22}\varepsilon_{12}) & \dots & E(\varepsilon_{22}\varepsilon_{1T}) & E(\varepsilon_{22}\varepsilon_{21}) & \dots & E(\varepsilon_{22}\varepsilon_{NT}) \\
\vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
E(\varepsilon_{2T}\varepsilon_{11}) & E(\varepsilon_{2T}\varepsilon_{12}) & \dots & E(\varepsilon_{2T}\varepsilon_{1T}) & E(\varepsilon_{2T}\varepsilon_{21}) & \dots & E(\varepsilon_{2T}\varepsilon_{NT}) \\
\vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
E(\varepsilon_{N1}\varepsilon_{11}) & E(\varepsilon_{N1}\varepsilon_{12}) & \dots & E(\varepsilon_{N1}\varepsilon_{1T}) & E(\varepsilon_{N1}\varepsilon_{21}) & \dots & E(\varepsilon_{N1}\varepsilon_{NT}) \\
\vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
E(\varepsilon_{NT}\varepsilon_{11}) & E(\varepsilon_{NT}\varepsilon_{12}) & \dots & E(\varepsilon_{NT}\varepsilon_{1T}) & E(\varepsilon_{NT}\varepsilon_{21}) & \dots & E(\varepsilon_{NT}^2)
\end{bmatrix}
$$

GLS estimation (and FGLS) relies on knowledge or estimation of the variance-covariance matrix V = E ($\varepsilon \varepsilon^T$)

$$\hat{\boldsymbol{\beta}}_{GLS} = \left(\boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{y}$$

$$\tilde{\boldsymbol{\beta}}_{FGLS} = \left(\boldsymbol{X}^T \hat{\boldsymbol{V}}^{-1} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \hat{\boldsymbol{V}}^{-1} \boldsymbol{y}$$

$$Est.\ Var-cov(\tilde{\boldsymbol{\beta}}_{FGLS}) = \left(X^T \hat{V}^{-1} X\right)^{-1}$$

and $\hat{V}$ is consistent estimate of V

Convenient decomposition of variance-covariance matrix for panel

$$E(\varepsilon\varepsilon^{T}) = \begin{bmatrix} \sigma_{11}\Omega_{11} & \sigma_{12}\Omega_{12} & \cdots & \sigma_{1N}\Omega_{1N} \\ \sigma_{21}\Omega_{21} & \sigma_{22}\Omega_{22} & \cdots & \sigma_{2N}\Omega_{2N} \\ \sigma_{31}\Omega_{31} & \sigma_{32}\Omega_{32} & \cdots & \sigma_{3N}\Omega_{3N} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{N1}\Omega_{N1} & \sigma_{N2}\Omega_{N2} & \cdots & \sigma_{NN}\Omega_{NN} \end{bmatrix}$$

where Ω_{ij} is a T x T matrix

A: suppose no correlation of error among individuals

1. Easiest case

Suppose no heteroskedasticity, cross-sectional independence and no autocorrelation

$$E(\varepsilon_{it}^2)=\sigma^2 \ , \quad E(\varepsilon_{it}\varepsilon_{jt})=0 \quad , \quad i \neq j \ , \quad E(\varepsilon_{it}\varepsilon_{i,t-1})=0$$

2. Group-wise heteroskedasticity

Suppose heteroskedasticity, cross-sectional independence and no autocorrelation

$$E(\varepsilon_{it}^2)=\sigma_i^2 \ , \quad E(\varepsilon_{it}\varepsilon_{jt})=0 \quad , \quad i \neq j \ , \quad E(\varepsilon_{it}\varepsilon_{i,t-1})=0$$

3. Autocorrelation

Suppose no heteroskedasticity, cross-sectional independence and autocorrelation

$$E(\varepsilon_{it}^2)=\sigma^2 \ , \quad E(\varepsilon_{it}\varepsilon_{jt})=0 \quad , \quad i \neq j \ , \quad \varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + u_{it}$$

Suppose no heteroskedasticity, cross-sectional independence and autocorrelation is the same for all individuals

$$E(\varepsilon_{it}^2)=\sigma^2 \ , \quad E(\varepsilon_{it}\varepsilon_{jt})=0 \quad , \quad i \neq j \ , \quad \varepsilon_{it} = \rho \varepsilon_{i,t-1} + u_{it}$$

4. Both

Suppose heteroskedasticity, cross-sectional independence and autocorrelation

$$E(\varepsilon_{it}^2)=\sigma_i^2 \ , \quad E(\varepsilon_{it}\varepsilon_{jt})=0 \quad , \quad i \neq j \ , \quad \varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + u_{it}$$

Assume further:

$u_{it} \sim N(0, \sigma^2_{ui})$ heteroskedastic .... and homoskedastic $N(0, \sigma^2)$

$$\varepsilon_{i1} \sim N\left(0\ ,\ \frac{\sigma^2_{ui}}{1-\rho_i^2}\right)$$ ..... For autocorrelation

1.

$\sigma_{ij} = 0$ for $i \neq j$ and $\sigma_{ii} = \sigma^2$, $\Omega_{ij} = 0$ for $i \neq j$ and $\Omega_{ii} = 1$,

$\boldsymbol{V} = \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega}$, $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_N$, $\boldsymbol{\Omega} = \boldsymbol{I}_T$ and therefore $\boldsymbol{V} = \sigma^2 \boldsymbol{I}_{NT}$,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$ Dimension is N x N

$$\boldsymbol{V} = \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$ Dimension is NT x NT

which means OLS estimation.

2.

$\sigma_{ij} = 0$ for $i \neq j$ and $\sigma_{ii} = \sigma_i^2$, $\Omega_{ij} = 0$ for $i \neq j$ and $\Omega_{ii} = 1$,

$$\boldsymbol{V} = \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega} \;, \quad \boldsymbol{\Omega} = \boldsymbol{I_T}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix} \qquad \boldsymbol{V} = \begin{bmatrix} \sigma_1^2[\boldsymbol{I_T}] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma_2^2[\boldsymbol{I_T}] & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_N^2[\boldsymbol{I_T}] \end{bmatrix}$$

N x N NT x NT

which means GLS estimation with group-wise heteroskedasticity

$$\hat{\sigma}_i^2 = \frac{e_i^T e_i}{T_i} \qquad \hat{\sigma}_i^2 = \frac{e_i^T e_i}{T_i - K}$$

3. Autocorrelation, same for all groups

$$\boldsymbol{V}=\boldsymbol{\Sigma}\otimes\boldsymbol{\Omega} \quad , \ \boldsymbol{\Sigma}=\sigma^2 \ \boldsymbol{I_N}$$

$$\boldsymbol{\Omega}=\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix}$$

If autocorrelation differs across groups:

$$\boldsymbol{V}=\sigma^2\begin{bmatrix} [\boldsymbol{\Omega_1}] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & [\boldsymbol{\Omega_2}] & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & [\boldsymbol{\Omega_N}] \end{bmatrix} \text{where } \boldsymbol{\Omega_i}=\begin{bmatrix} 1 & \rho_i & \rho_i^2 & \cdots & \rho_i^{T-1} \\ \rho_i & 1 & \rho_i & \cdots & \rho_i^{T-2} \\ \rho_i^2 & \rho_i & 1 & \cdots & \rho_i^{T-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 \end{bmatrix}$$

From small samples use sample correlation

$$\hat{\rho}_i = \frac{\sum_{t=2}^{T_i} e_{it} e_{i,t-1}}{\sum_{t=2}^{T_i} e_{i,t-1}^2}$$

$$\hat{\rho}_i = \frac{\sum_{t=2}^{T_i} e_{it} e_{i,t-1}}{\sqrt{\sum_{t=2}^{T_i} e_{i,t-1}^2} \sqrt{\sum_{t=2}^{T_i} e_{it}^2}}$$

$$\hat{\rho} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T_i} e_{it} e_{i,t-1}}{\sum_{i=1}^{N} \sum_{t=2}^{T_i} e_{i,t-1}^2}$$

$$\hat{\rho} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T_i} e_{it} e_{i,t-1}}{\sqrt{\sum_{i=1}^{N} \sum_{t=2}^{T_i} e_{i,t-1}^2} \sqrt{\sum_{i=1}^{N} \sum_{t=2}^{T_i} e_{it}^2}}$$

and then use Prais-Winsten or Cochran-Orcutt transformation:

$$W_{i1}^* = \sqrt{1-\hat{\rho}_i^2}\; W_{i1}$$ Prais-Winsten

$$W_{it}^* = W_{it} - \hat{\rho}_i \; W_{i,t-1}$$ Cochran-Orcutt and P-W

W are variables Y and all K regressors (X)

4. Autocorrelation and heteroskedasticity

$$\boldsymbol{V}=\begin{bmatrix} \sigma_1^2[\boldsymbol{\Omega_1}] & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_2^2[\boldsymbol{\Omega_2}] & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \sigma_N^2[\boldsymbol{\Omega_N}] \end{bmatrix}, \quad \boldsymbol{\Omega_i}=\begin{bmatrix} 1 & \rho_i & \rho_i^2 & \cdots & \rho_i^{T-1} \\ \rho_i & 1 & \rho_i & \cdots & \rho_i^{T-2} \\ \rho_i^2 & \rho_i & 1 & \cdots & \rho_i^{T-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 \end{bmatrix}$$

Procedure:

Apply Cochran-Orcutt, then the resulting residuals should be estimates of u_{it}^*, denote them $\hat{u}_{it}^*$, which are heteroskedastic across groups but not autocorrelated. We can use them to estimate the (group) variance of u_{it}, i.e. σ^2_{ui}

$$\hat{\sigma}_{ui}^2 = s_{ui}^2 = \frac{1}{T_i - K} \sum_{t=1}^{T_i} \hat{u}_{it}^{*2} \qquad\qquad \hat{\sigma}_{ui}^2 = s_{ui}^2 = \frac{1}{T_i} \sum_{t=1}^{T_i} \hat{u}_{it}^{*2}$$

and then use heteroskedastic approach:

$$W_{it}^{**} = \frac{W_{it}^*}{s_{ui}}$$

W^* are transformed variables Y^* and all K regressors (X^*)

Note: $var(e_i) = s_i^2 = \dfrac{s_{ui}^2}{1-\hat{\rho}_i^2}$

B: suppose correlation of error among individuals

Suppose also heteroskedasticity and autocorrelation

$E(\varepsilon_{it}^2)=\sigma_{ii}$ heteroskedasticity

$E(\varepsilon_{it}\varepsilon_{jt})=\sigma_{ij} \quad , \quad i \neq j$ mutual correlation between errors

$\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + u_{it}$ Autocorrelation structure

$u_{it} \sim N(0 \ , \ \varphi_{ii})$ Innovative error

$E(\varepsilon_{i,t-1} u_{jt})=0$ Innovative error is independent on past

$E(u_{it} u_{jt})=\varphi_{ij}$ This is the mutual correlation of innovative part

$E(u_{it} u_{js})=0 \ , \quad t \neq s$ Innovative error is not autoregressive

$\sigma_{ii}=\dfrac{\varphi_{ii}}{1-\rho_i^2} \quad \sigma_{ij}=\dfrac{\varphi_{ij}}{1-\rho_i\rho_j}$ Variance-covariance terms

Assumption on first errors:

$$\varepsilon_{i1} \sim N\left(0 \ , \ \frac{\varphi_{ii}}{1-\rho_i^2}\right) \qquad E(\varepsilon_{i1}\varepsilon_{j1}) = \frac{\varphi_{ij}}{1-\rho_i\rho_j}$$

We can then write down the variance-covariance matrix:

$$\boldsymbol{V}=\begin{bmatrix} \sigma_{11}[\boldsymbol{\Omega_{11}}] & \sigma_{12}[\boldsymbol{\Omega_{12}}] & \cdots & \sigma_{1N}[\boldsymbol{\Omega_{1N}}] \\ \sigma_{21}[\boldsymbol{\Omega_{21}}] & \sigma_{22}[\boldsymbol{\Omega_{22}}] & \cdots & \sigma_{2N}[\boldsymbol{\Omega_{2N}}] \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_{N1}[\boldsymbol{\Omega_{N1}}] & \sigma_{N2}[\boldsymbol{\Omega_{N2}}] & \cdots & \sigma_{NN}[\boldsymbol{\Omega_{NN}}] \end{bmatrix}$$

$$\boldsymbol{\Omega_{ij}}=\begin{bmatrix} 1 & \rho_j & \rho_j^2 & \cdots & \rho_j^{T-1} \\ \rho_i & 1 & \rho_j & \cdots & \rho_j^{T-2} \\ \rho_i^2 & \rho_i & 1 & \cdots & \rho_j^{T-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 \end{bmatrix}$$

We can use similar procedure as before:

1. Estimate autocorrelation

$$\hat{\rho}_i = \frac{\sum_{t=2}^{T_i} e_{it} e_{i,t-1}}{\sum_{t=2}^{T_i} e_{i,t-1}^2}$$

And transform:

$$W_{i1}^* = \sqrt{1-\hat{\rho}_i^2}\; W_{i1}$$ Prais-Winsten

$$W_{it}^* = W_{it} - \hat{\rho}_i\; W_{i,t-1}$$ Cochran-Orcutt and P-W

W are variables Y and all K regressors (X)

Apply again OLS, get estimates of u_{it}^*, denoted $\hat{u}_{it}^*$, and use them to estimate variance-covariances:

$$s_{ij} = \frac{\hat{\varphi}_{ij}}{1-\rho_i\rho_j} \quad \text{and} \quad \hat{\varphi}_{ij} = \frac{1}{T-K}\sum_{t=1}^{T}\hat{u}_{it}^*\hat{u}_{jt}^*$$

Note: form unbalanced panels you can use

$$\hat{\varphi}_{ij} = \frac{1}{T}\sum_{overlapped \quad t}\hat{u}_{it}^*\hat{u}_{jt}^*, \qquad T = |overlapped \quad t|$$

$$\tilde{\boldsymbol{\beta}}_{FGLS} = \left(\boldsymbol{X}^{*T}\hat{\boldsymbol{\Phi}}^{-1}\boldsymbol{X}^*\right)^{-1}\boldsymbol{X}^{*T}\hat{\boldsymbol{\Phi}}^{-1}\boldsymbol{y}^*$$

$$Est.Var-cov(\tilde{\boldsymbol{\beta}}_{FGLS}) = \left(\boldsymbol{X}^{*T}\hat{\boldsymbol{\Phi}}^{-1}\boldsymbol{X}^*\right)^{-1}$$

$$\hat{\boldsymbol{\Phi}} = \begin{bmatrix} \hat{\varphi}_{11}[\boldsymbol{I_T}] & \hat{\varphi}_{12}[\boldsymbol{I_T}] & \cdots & \hat{\varphi}_{1N}[\boldsymbol{I_T}] \\ \hat{\varphi}_{21}[\boldsymbol{I_{21}}] & \hat{\varphi}_{22}[\boldsymbol{I_T}] & \cdots & \hat{\varphi}_{2N}[\boldsymbol{I_T}] \\ \vdots & \vdots & \cdots & \vdots \\ \hat{\varphi}_{N1}[\boldsymbol{I_T}] & \hat{\varphi}_{N2}[\boldsymbol{I_T}] & \cdots & \hat{\varphi}_{NN}[\boldsymbol{I_T}] \end{bmatrix}$$

Note: The upper left term of inverse of Ω_{ij}, Ω_{ij}^{-1}, is different from term of matrix $(P_i^T P_j)^T$ (apart form scaling factor between matrices):

$$1 \neq \sqrt{1-\rho_i^2}\sqrt{1-\rho_j^2} + \rho_i\rho_j$$

Solution: use C-O method, or autocorrelation coefficients should be similar. Discrepancy is asymptotically irrelevant anyway.

Testing

Group-wise heteroskedasticity

$s_i = s$ for all i

$$LM = W = \frac{T}{2} \sum_{i=1}^{N} \left[\frac{s_i^2}{s^2} - 1 \right]^2 \sim \chi^2_{N-1}$$

$$LR = NT \ln \hat{\sigma}^2 - \sum_{i=1}^{N} T \sigma_i^2 \sim \chi^2_{N-1} \qquad \hat{\sigma}_i^2 = \frac{e_i^T e_i}{T_i} \quad , \quad \hat{\sigma}^2 = \frac{e^T e}{NT}$$

Group-wise correlation (with heteroskedasticity, unbalanced...)

$r_{ij}=0$ for all i, j, i<>j

$$LM = T \sum_{i=2}^{N} \sum_{j=1}^{i} r_{ij}^2 \sim \chi^2_{\frac{N(N-1)}{2}}$$

r^2 is correlation coeff. of resid.

$$LR = T \left(\sum_{i=1}^{N} \ln s_i^2 - \ln \det(\hat{\Sigma}) \right) \sim \chi^2_{\frac{N(N-1)}{2}}$$