Panel data

Jaroslav Dufek

Obsah

Model

$$y_{i,t} = \alpha + \mathbb{X}_{it}^T \beta + u_{it}$$

- $i \ldots$ panels (firms, countries, individual parameters)
- $t \ldots$ time
- $\alpha \ldots$ scalar
- $\beta \ldots$ regressors dim $K + 1$
- $\mathbb{X}_{it}^T \ldots$ observation for i-th panel and time t

One way error component model

$$u_{it} = \mu_i + \nu_{it} \tag{1}$$

- $\mu_i \ldots$ unobservable individual effects
- $\nu_{it} \ldots$ reminder disturbance

Vector notation

$$\mathbf{y} = \alpha \mathbf{1}_{NT} + \mathbb{X}\beta + \mathbf{u} = \mathbb{Z}\delta + \mathbf{u} \tag{2}$$

- $\mathbf{y} \ldots (NT \times 1)$
- $\mathbb{X} \ldots (NT \times K)$
- $\mathbb{Z} = (\mathbf{1}_{NT}, \mathbb{X})$
- $\delta^T = (\alpha, \beta^{\mathbf{T}})$
- $\mathbf{1}_{NT}$ vector of ones of dimension NT

Vector notation of error I

Then (1) is possible rewrite

$$\mathbf{u} = \mathbb{Z}_\mu \mu + \nu \tag{3}$$

- $\mathbf{u}^T = (u_{11}, \ldots, u_{1T}, u_{21}, \ldots, u_{2T}, \ldots, u_{N1}, \ldots, u_{NT})$
- $\mathbb{Z}_\mu = \mathbb{I}_N \otimes \mathbb{J}_T$
- $\mathbb{I}_N \ldots$ identity matrix of dimension N
- $A \otimes B \ldots$ Kronecker product
- $\mu^T = (\mu_1, \ldots, \mu_N)$; $\nu^T = (\nu_{11}, \ldots, \nu_{1T}, \ldots, \nu_{N1}, \ldots, \nu_{NT})$

$\Longrightarrow \mathbb{Z}_\mu \mathbb{Z}_\mu^T = \frac{1}{T} \mathbb{I}_N \mathbb{J}_T$

- $\mathbb{J}_T \ldots$ matrix of ones of dimension T

Vector notation of error II

Let's define

- $\mathbb{P} = \frac{1}{T}\mathbb{Z}_\mu(\mathbb{Z}_\mu^T\mathbb{Z}_\mu)^{-1}\mathbb{Z}_\mu^T \ldots$ projection matrix of $\mathbb{Z}_\mu$
- $\mathbb{Q} = \mathbf{1}_{NT} - \mathbb{P} \ldots$ deviation form individual means

$\Longrightarrow \mathbb{P}, \mathbb{Q}$ are idempotent matrices

- $\mathbb{P}^T = \mathbb{P}$
- $\mathbb{P}^2 = \mathbb{P}$
- rank($\mathbb{P}$) =tr($\mathbb{P}$) $= N$; rank($\mathbb{Q}$) =tr($\mathbb{Q}$) $= N(T-1)$

$\Longrightarrow \mathbb{P}, \mathbb{Q}$ are orthogonal i.e. $\mathbb{P}\mathbb{Q} = 0 \Longrightarrow \mathbb{P} + \mathbb{Q} = \mathbb{I}_{NT}$

The fixed effects models I

From (1) and (2)

$$\mathbf{y} = \alpha \mathbf{1}_{NT} + \mathbb{X}\beta + \mathbb{Z}_\mu \mu + \nu \tag{4}$$

- μ_i is fixed parameter for estimate
- ν_{it} iid; $E\nu_{it} = 0,\ var\nu_{it} = \sigma_\nu^2$
- $\forall i, t\ \mathbb{X}_{it},\ \nu_{it}$ are mutually independent

The fixed effects models II

From (4) we can easily obtain

$$\mathbf{y} = \mathbb{Z}\delta + \mathbb{Z}_\mu \mu + \nu \quad (5)$$

$\rightsquigarrow$ run the OLS for estimation of parameters α, β, μ

- $\mathbb{Z}$ is $NT \times (K+1)$
- $\mathbb{Z}_\mu$ is $NT \times N$

$\Longrightarrow$ for N huge the OLS needs to invert matrix of dimension $(N+K)$!!!

The fixed effects models III - Trick

We are interested in parameters α, β
$\rightsquigarrow$ LSDV (least squares dummy variables)
pre-multiplying (5) by matrix $\mathbb{Q}$

$$\mathbb{Q}\mathbf{y} = \mathbb{Q}\mathbb{X}\beta + \mathbb{Q}\nu$$

$\Longrightarrow$ typical regression $\tilde{\mathbf{y}} = \mathbb{Q}\mathbf{y}$ with typical representants $(y_{it} - \bar{y}_{i\cdot})$ on $\tilde{\mathbb{X}} = \mathbb{Q}\mathbb{X}$ with typical representants $(X_{it,k} - \bar{X}_{i\cdot,k})$ for k-th regressor where $k = 1, 2, \ldots, K$
$\rightsquigarrow$ inversion of matrix of type $(K \times K)$

Estimation by the OLS method

- $\tilde{\beta} = (\mathbb{X}^T\mathbb{Q}\mathbb{X})^{-1}\mathbb{X}^T\mathbb{Q}\mathbf{y}$
- $\text{var}\tilde{\beta} = \sigma_\nu^2(\mathbb{X}^T\mathbb{Q}\mathbb{X})^{-1} = \sigma_\nu^2(\tilde{\mathbb{X}}^T\tilde{\mathbb{X}})^{-1}$

The fixed effects models IV - Remark

- Simple regression

$$y_{it} = \alpha + \beta X_{it} + \mu_i + \nu_{it} \qquad (*)$$

- Averaging

$$\bar{y}_{i.} = \alpha + \beta \bar{X}_{i.} + \mu_i + \bar{\nu}_{i.} \qquad (**)$$

- $(*) - (**)$

$$y_{it} - \bar{y}_{i.} = \beta(X_{it} - \bar{X}_{i.}) + (\nu_{it} - \bar{\nu}_{i.}) \tag{6}$$

- Averaging again

$$\bar{y}_{..} = \alpha + \beta \bar{X}_{..} + \bar{\nu}_{..}$$

- subject to $\sum_{i=1}^{N} \mu_i = 0 \ldots$ preventing from dummy trap
- it is possible to estimate only parameters β and $(\alpha + \mu_i)$
- impossible of separation of parameters α and μ_i

The fixed effects models V - Significance

Testing of significance of fixed effects

- $H_0: \ \mu_1 = \mu_2 = \cdots = \mu_{N-1} = 0$
- testing statistics under H_0

$$F_0 = \frac{(RRSS - URSS)/(N-1)}{URSS/(NT - N - K)} \overset{H_0}{\sim} F_{N-1, N(T-1)-K}$$

- RRSS resticted residual sum of squares
- URSS unresticted residual sum of squares

The fixed effects models VI - Robust estimates of std errors

- equation for regression after pre-multiplying by $\mathbb{Q}$

$$\tilde{\mathbf{y}}_i = \tilde{\mathbb{X}}_i \beta + \tilde{\nu}_i$$

- robust least squares

$$N^{1/2}(\tilde{\beta} - \beta) \sim N(0, M^{-1} V M^{-1})$$

- $M = plim(\bar{\mathbb{X}}\mathbb{X})/N; \;\; V = plim \sum_1^N (\tilde{\mathbb{X}}_i^T \mathbf{\Omega}_i \mathbb{X}_i)/N$
- $\mathbf{\Omega}_i = E(\nu_i \nu_i^T)$

$$\rightsquigarrow \text{var}(\tilde{\beta}) = (\tilde{\mathbb{X}}^T \tilde{\mathbb{X}})^{-1} \left[\sum_{i=1}^{N} \tilde{\mathbb{X}}^T \tilde{\mathbf{u}}_i \tilde{\mathbf{u}}_i^T \tilde{\mathbb{X}}_i \right] (\tilde{\mathbb{X}}^T \tilde{\mathbb{X}})^{-1}$$

- $\tilde{\mathbf{u}}_i = \tilde{\mathbf{y}}_i - \tilde{\mathbb{X}}_i \tilde{\beta}$

The random effects model I

$$\mathbf{y} = \alpha \mathbf{1}_{NT} + \mathbb{X}\beta + \mathbb{Z}_\mu \mu + \nu$$

- $\mu_i \ldots$ iid, $E\mu_i = 0$, $var\mu_i = \sigma_\mu^2$
- $\nu_{it} \ldots$ iid, $E\nu_{it} = 0$, $var\nu_{it} = \sigma_\nu^2$
- μ_i, ν_{it}, X_{it} mutually independent
- from (3) $u = \mathbb{Z}_\mu \mu + \nu$ we can obtain correlation matrix

$$\mathbf{\Omega} = E(\mathbf{u}\mathbf{u}^T) = \mathbb{Z}_\mu E(\mu\mu^T)\mathbb{Z}_\mu^T + E(\nu\nu^T) = \sigma_\mu^2(\mathbb{I}_N \otimes \mathbb{J}_T) + \sigma_\nu^2(\mathbb{I}_N \otimes \mathbb{J}_T) \quad (*)$$

$$cov(u_{it}, u_{js}) = \begin{cases} \sigma_\mu^2 + \sigma_\nu^2 & \text{for } i = j, t = s \\ \sigma_\mu^2 & \text{for } i = j, t \neq s \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

The random effects model II - GLS Estimator

- For GLS estimator is needed $\boldsymbol{\Omega}^{-1}$ that is matrix of type $(NT \times NT)$

$\rightsquigarrow$ **Trick** Wansbeek(1982,1983) for $\boldsymbol{\Omega}^{-1}, \ \boldsymbol{\Omega}^{-1/2}$

- For substitution $\mathbb{J}_T$ by $\bar{\mathbb{J}}_T$ and $\mathbb{I}_T$ by $(\mathbb{E}_T + \bar{\mathbb{J}}_T); \mathbb{E}_T = \mathbb{I}_T - \bar{\mathbb{J}}_T$ in $(*)$ obtain

$$\boldsymbol{\Omega} = T\sigma_\mu^2(\mathbb{I}_N \otimes \bar{\mathbb{J}}_T) + \sigma_\nu^2(\mathbb{I}_N \otimes \mathbb{E}_T) + \sigma_\nu^2(\mathbb{I}_N \otimes \bar{\mathbb{J}}_T)$$

i.e.

$$\boldsymbol{\Omega} = (T\sigma_\mu^2 + \sigma_\nu^2)(\mathbb{I}_N \otimes \bar{\mathbb{J}}_T) + \sigma_\nu^2(\mathbb{I}_N \otimes \mathbb{E}_T) = \sigma_1^2\mathbb{P} + \sigma_\nu^2\mathbb{Q}, \qquad (8)$$

kde $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$

The random effects model III - GLS Estimator

- (8) is spectral decomposition of Ω where first unique characteristic root σ_1^2 of multiplicity N and second root σ_ν^2 of multiplicity $N(T-1)$
- From properties of $\mathbb{P}, \mathbb{Q}$
  - $\boldsymbol{\Omega}^{-1} = \frac{1}{\sigma_1^2}\mathbb{P} + \frac{1}{\sigma_\nu^2}\mathbb{Q}$
  - $\boldsymbol{\Omega}^{-1/2} = \frac{1}{\sigma_1}\mathbb{P} + \frac{1}{\sigma_\nu}\mathbb{Q}$
- generally holds $\boldsymbol{\Omega}^r = \left(\sigma_1^2\right)^r \mathbb{P} + \left(\sigma_\nu^2\right)^r \mathbb{Q}$
- after pre-multiplying of (2) by $\sigma_\nu \boldsymbol{\Omega}^{-1/2} = \mathbb{Q} + \frac{\sigma_\nu}{\sigma_1}\mathbb{P}$ we obtain regression $\mathbf{y}^* = \sigma_\nu \boldsymbol{\Omega}^{-1/2}\mathbf{y}$
  $\rightsquigarrow$ OLS - inversion of matrix of dimension $(K+1)$

The best quadratic unbiased (BQU) estimator

- BQU estimator comes form spectral decomposition $\mathbf{\Omega}$

$$\hat{\sigma}_1^2 = \frac{\mathbf{u}^T \mathbb{P} \mathbf{u}}{tr(\mathbb{P})} = \frac{T \sum_{i=1}^{N} \bar{u}_{i\cdot}^2}{N} \tag{9}$$

$$\hat{\sigma}_\nu^2 = \frac{\mathbf{u}^T \mathbb{Q} \mathbf{u}}{tr(\mathbb{Q})} = \frac{\sum_{i=1}^{N} \sum_{t+1}^{T} (u_{it} - \bar{u}_{i\cdot})^2}{N(T-1)} \tag{10}$$

Least squares dummies variables (LSDV)

- OLS estimates are still unbiased and consistent but not efficient $\rightsquigarrow$ use LSDV residuals instead of OLS residuals
- $\tilde{\mathbf{u}} = \mathbf{y} - \tilde{\alpha}\mathbf{1}_{NT} - \mathbb{X}\tilde{\beta}$,
  - $\tilde{\alpha} = \bar{y}_{..} - \bar{\mathbb{X}}_{..}^T \tilde{\beta}$
  - $\bar{\mathbb{X}}_{..}^T \ldots (1 \times K)$ vector of averages of all regressors

$$\begin{pmatrix} \sqrt{NT}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) \\ \sqrt{N}(\hat{\sigma}_\mu^2 - \sigma_\mu^2) \end{pmatrix} \sim N\left(0, \begin{pmatrix} 2\sigma_\nu^4 & 0 \\ 0 & 2\sigma_\mu^4 \end{pmatrix} \right),$$

where $\hat{\sigma}_\mu^2 = \frac{(\hat{\sigma}_1^2 - \hat{\sigma}_\nu^2)}{T}$

Double regression I

Swamy (1972)

- first regression equation
  - form (6) $y_{it} - \bar{y}_{i\cdot} = \beta(X_{it} - \bar{X}_{i\cdot}) + (\nu_{it} - \bar{\nu}_{i\cdot})$

$$\hat{\hat{\sigma}}_{\nu}^{2} = \frac{\left[\mathbf{y}^T \mathbb{Q}\mathbf{y} - \mathbf{y}^T \mathbb{Q}\mathbb{X}(\mathbb{X}^T \mathbb{Q}\mathbb{X})^{-1}\mathbb{X}^T \mathbb{Q}\mathbf{y}\right]}{N(T-1) - K}$$

- second regression equation (cross-section regression)

$$\sqrt{T}\bar{y}_i = \sqrt{T}\alpha + \sqrt{T}\bar{\mathbb{X}}_i^T \beta + \sqrt{T}\bar{\mathbf{u}}_i,$$

  where $\mathrm{var}(\sqrt{T}\bar{\mathbf{u}}_{i\cdot}) = \sigma_1^2$

  - $$\hat{\hat{\sigma}}_1^2 = \frac{\mathbf{y}^T \mathbb{P}\mathbf{y} - \mathbf{y}^T \mathbb{P}\mathbb{Z}(\mathbb{Z}^T \mathbb{P}\mathbb{Z})^{-1}\mathbb{Z}^T \mathbb{P}\mathbf{y}}{N - K - 1}$$

Double regression II - Remark

- The double regression can be rewritten in matrices

$$\begin{pmatrix} \mathbb{Q}\mathbf{y} \\ \mathbb{P}\mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbb{Q}\mathbb{Z} \\ \mathbb{P}\mathbb{Z} \end{pmatrix} \delta + \begin{pmatrix} \mathbb{Q}\mathbf{u} \\ \mathbb{P}\mathbf{u} \end{pmatrix}$$

- transformed error has zero mean and covariance matrix

$$\begin{pmatrix} \sigma_\nu^2 \mathbb{Q} & 0 \\ 0 & \sigma_1^2 \mathbb{P} \end{pmatrix}$$

GLS

- if we want to get rid of constant α

$$\begin{pmatrix} \mathbb{Q}\mathbf{y} \\ (\mathbb{P} - \bar{\mathbb{J}}_{NT})\mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbb{Q}\mathbb{X} \\ (\mathbb{P} - \bar{\mathbb{J}}_{NT})\mathbb{X} \end{pmatrix} \beta + \begin{pmatrix} \mathbb{Q}\mathbf{u} \\ (\mathbb{P} - \bar{\mathbb{J}}_{NT})\mathbf{u} \end{pmatrix} \quad (11)$$

- we obtain it by using facts that $\mathbb{Q}\mathbf{1}_{NT} = 0$ and $(\mathbb{P} - \bar{\mathbb{J}}_{NT})\mathbf{1}_{NT} = 0$
- transformed error has zero mean and covariance matrix

$$\begin{pmatrix} \sigma_\nu^2 \mathbb{Q} & 0 \\ 0 & \sigma_1^2 \mathbb{P} \end{pmatrix}$$

GLS II

GLS on (11):

$$\hat{\beta}_{GLS} = \left[\frac{\mathbb{X}^T\mathbb{Q}\mathbb{X}}{\sigma_\nu^2} + \frac{\mathbb{X}^T(\mathbb{P}-\bar{\mathbb{J}}_{NT})\mathbb{X}}{\sigma_1^2}\right]^{-1} \times \left[\frac{\mathbb{X}^T\mathbb{Q}\mathbf{y}}{\sigma_\nu^2} + \frac{\mathbb{X}^T(\mathbb{P}-\bar{\mathbb{J}}_{NT})\mathbf{y}}{\sigma_1^2}\right] =$$

$$= \left[\mathbb{W}_{XX} + \Phi^2\mathbb{B}_{XX}\right]^{-1}\left[\mathbb{W}_{Xy} + \Phi^2\mathbb{B}_{Xy}\right]$$

- $\mathrm{var}\hat{\beta}_{GLS} = \sigma_\nu^2 = [\mathbb{W}_{XX} + \Phi^2\mathbb{B}_{XX}]$
- $\mathbb{W}_{XX} = \mathbb{X}^T\mathbb{Q}\mathbb{X};\ \mathbb{B}_{XX} = \mathbb{X}^T(\mathbb{P}-\bar{\mathbb{J}}_{NT})\mathbb{X}; \Phi^2 = \frac{\sigma_\nu^2}{\sigma_1^2}$

GLS III

- $\tilde{\beta}_W = \mathbb{W}_{XX}^{-1}\mathbb{W}_{Xy}$
- $\tilde{\beta}_B = \mathbb{B}_{XX}^{-1}\mathbb{B}_{Xy}$

$\Rightarrow$

$$\hat{\beta}_{GLS} = \mathbb{W}_1\tilde{\beta}_W + \mathbb{W}_2\tilde{\beta}_B$$

- $\mathbb{W}_1 = (\mathbb{W}_{XX} + \Phi^2\mathbb{B}_{XX})^{-1}\mathbb{W}_{XX}$
- $\mathbb{W}_2 = (\mathbb{W}_{XX} + \Phi^2\mathbb{B}_{XX})^{-1}(\Phi^2\mathbb{B}_{XX}) = \mathbb{I} - \mathbb{W}_1$

$\Rightarrow$ $\hat{\beta}_{GLS}$ is weighted estimate

Maximum likelihood estimation

- under normality of disturbances

$$L(\alpha, \beta, \Phi^2, \sigma_{nu}^2) = \text{const.} - \frac{NT}{2} \log \sigma_\nu^2 + \frac{N}{2} \log \Phi^2 - \frac{1}{2\sigma_\nu^2} \mathbf{u}^T \Sigma^{-1} \mathbf{u}$$

  - $\Sigma = \mathbb{Q} + \Phi^{-2}\mathbb{P}$
- $\rightsquigarrow$ $\hat{\beta}_{\text{mle}}$; $\hat{\alpha}_{\text{mle}} = \bar{y}_{..} - \mathbb{X}_{..}^T \hat{\beta}_{\text{mle}}$
- $\hat{\sigma}^2_{\text{mle}} = \frac{\mathbf{d}^T [\mathbb{Q} + \Phi^2 (\mathbb{P} - \bar{\mathbb{J}}_{NT})] \mathbf{d}}{NT}$
  - $\mathbf{d} = \mathbf{y} - \mathbb{X}\hat{\beta}_{\text{mle}}$

Concentrated likelihood

$$L_C(\beta, \Phi^2) = \text{const.} - \frac{NT}{2} \log \left\{ \mathbf{d}^T [\mathbb{Q} + \Phi^2(\mathbb{P} - \bar{\mathbb{J}}_{NT})] \mathbf{d} \right\} + \frac{N}{2} \log \Phi^2 \quad (12)$$

- max (12) over β

$$\hat{\beta}_{\text{mle}} = \left[\mathbb{X}^T (\mathbb{Q} + \Phi^2(\mathbb{P} - \bar{\mathbb{J}}_{NT})) \mathbb{X} \right]^{-1} \mathbb{X}^T \left[\mathbb{Q} + \Phi^2(\mathbb{P} - \bar{\mathbb{J}}_{NT}) \right] \mathbf{y}$$

Prediction

- predict S period ahead for i-th individual

$$\hat{\mathbf{y}}_{i,T+s} = \mathbb{Z}_{i,T+s}^T \hat{\delta}_{\mathrm{GLS}} + \mathbf{w}^T \Omega^{-1} \hat{\mathbf{u}}_{\mathrm{GLS}} \quad s \geq 1$$

- $\hat{\mathbf{u}}_{\mathrm{GLS}} = \mathbf{y} - \mathbb{Z}\hat{\delta}_{\mathrm{GLS}}$
- $\mathbf{w} = E[u_{i,T+s} \cdot u]$
- $u_{i,T+s} = \mu_i + \nu_{i,T+s}$

Literatura

Badi H. Baltagi:
Econometrics Analysis of Panel Data
JOHN WILEY & SONS, LTD, 2001.

Swamy:
The exact finite sample properties of the estimators of coefficients in the error components regression models.
Econometrica 40: 261–275, 1972.

Wansbeek:
A simple way to obtain the spectral decomposition of variance components models for balanced data.
Communications in Statistics A11: 2105–2112, 1982.

Wansbeek:
A note on spectral decomposition and maximum likelihood estimation of ANOVA models with balanced data.
Statistics and Probability Letters 1: 213–215, 1983.

Děkuji za pozornost