One-way error component models. Fixed and random effects model

Jakub Mućk
SGH Warsaw School of Economics

One-way error component model

- In the model:

$$y_{it} = \alpha + X_{it}'\beta + u_{it} \quad i \in \{1, \dots, N\}, t \in \{1, \dots, T\} \tag{1}$$

it is assumed that all units are homogeneous. Why?

- **One-way error component model:**

$$u_{i,t} = \mu_i + \varepsilon_{i,t} \tag{2}$$

where:
  - μ_i – **the unobservable individual-specific effect**;
  - $\varepsilon_{i,t}$ – the remainder disturbance.

- In the model:

$$y_{it} = \alpha + X_{it}'\beta + u_{it} \quad i \in \{1,\ldots,N\}, t \in \{1,\ldots,T\} \tag{1}$$

it is assumed that all units are homogeneous. Why?

- **One-way error component model:**

$$u_{i,t} = \mu_i + \varepsilon_{i,t} \tag{2}$$

where:

  - μ_i – **the unobservable individual-specific effect**;
  - $\varepsilon_{i,t}$ – the remainder disturbance.

- **The most popular estimators:**
  - **The fixed effects (FE) estimator.**
  - **The random effects (RE) estimator.**

Fixed effects model

- We can relax assumption that all individuals have the same coefficients

$$y_{it} = \alpha_i + \beta_1 x_{1it} + \ldots + \beta_k x_{kit} + u_{it} \tag{3}$$

- Note that an i subscript is added to only intercept α_i but the slope coefficients, β_1, . . ., β_k are constant for all individuals.
- **An individual intercept (α_i)** are include to **control** for individual-specific and **time-invariant characteristics. That intercepts are called fixed effects**.
- Fixed effects capture the **individual heterogeneity.**
- The estimation:
  i) The least squares dummy variable estimator
  ii) The fixed effects/within estimator

- The natural way to estimate fixed effect for all individuals is to include an indicator variable. For example, for the first unit:

$$\mathcal{D}_{1i} = \begin{cases} 1 & i = 1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

  The number of dummy variables equals N. **It is not feasible to use the least square dummy variable estimator when N is large**

- We might rewrite the fixed regression as follows:

$$y_{it} = \sum_{j=1}^{N} \alpha_i \mathcal{D}_{ji} + \beta_1 x_{1it} + \ldots + \beta_k x_{kit} + u_{i,t} \tag{5}$$

- Why α is missing?

- Let us start with simple fixed effects specification for individual i:

$$y_{it} = \alpha_i + \beta_1 x_{1it} + \ldots + \beta_k x_{kit} + u_{it} \qquad t = 1, \ldots, T \tag{6}$$

- **Average the observation across time** and using the assumption on time-invariant parameters we get:

$$\bar{y}_i = \alpha_i + \beta_1 \bar{x}_{1i} + \ldots + \beta_k \bar{x}_{ki} + \bar{u}_i \tag{7}$$

where $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$, $\bar{x}_{1i} = \frac{1}{T}\sum_{t=1}^{T} x_{1it}$, $\bar{x}_{ki} = \frac{1}{T}\sum_{t=1}^{T} x_{kit}$ and $\bar{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_i$

- Now we substract (7) from (6)

$$(y_{it} - \bar{y}_i) = \underbrace{(\alpha_i - \alpha_i)}_{=0} + \beta_1(x_{1it} - \bar{x}_{1i}) + \ldots + \beta_k(x_{itk} - \bar{x}_{ki}) + (u_{it} - \bar{u}_i) \tag{8}$$

Using notation: $\tilde{y}_{it} = (y_{it} - \bar{y}_i)$, $\tilde{x}_{1it} = (x_{1it} - \bar{x}_{1i})$, $\tilde{x}_{kit} = (x_{kit} - \bar{x}_{ki})$ $\tilde{u}_{it} = (u_{it} - \bar{u}_i)$,
we get

$$\tilde{y}_{it} = \beta_1 \tilde{x}_{1t} + \ldots + \beta_k \tilde{x}_{kt} + \tilde{u}_{i,t} \tag{9}$$

Note that we **do not estimate directly fixed effects**.

Poolability test

- We can test estimates of intercept to verify whether the fixed effects are different among units:

$$\mathcal{H}_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_N. \quad (10)$$

- The poolability test is design to test joint significance of individual-specific intercepts.
- To test (10) we estimate: i) **unrestricted model** (the least squares dummy variable estimator) and ii) **restricted model** (pooled regression). Then we calculate sum of squared errors for both models: SSE_U and SSE_R.

$$\mathcal{F} = \frac{(SSE_R - SSE_U)/(N-1)}{SSE_U/(NT-K)} \quad (11)$$

if null is true then $\mathcal{F} \sim \mathcal{F}_{(N-1,NT-K)}$.

- The within transformation, by construction, eliminates all time invariant explanatory variables.
- A common practice is to use robust/clustered standard errors.
  - In the FE estimation, the heterogeneity in outcome is only taken into account. However, there could be systematic differences in variation of outcome for each individual.
- In the FE estimation, individual-specific effects could be potentially correlated with explanatory variables since they are fixed.

Random Effect model

- Let's assume the following model

$$y_{it} = \alpha + X_{it}'\beta + u_{it}, \quad i \in \{1, \ldots, N\}, t \in \{1, \ldots, T\}. \tag{12}$$

- The error component (u_t) is the sum of the individual specific random component (μ_i) and idiosyncratic disturbance $(\varepsilon_{i,t})$:

$$u_{it} = \mu_i + \varepsilon_{it}, \tag{13}$$

where $\mu_i \sim \mathcal{N}(0, \sigma_\mu^2)$;
and $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

- Let's assume the following model

$$y_{it} = \alpha + X_{it}'\beta + u_{it}, \quad i \in \{1, \dots, N\}, t \in \{1, \dots, T\}. \tag{12}$$

- The error component (u_t) is the sum of the individual specific random component (μ_i) and idiosyncratic disturbance $(\varepsilon_{i,t})$:

$$u_{it} = \mu_i + \varepsilon_{it}, \tag{13}$$

where $\mu_i \sim \mathcal{N}(0, \sigma_\mu^2)$;
and $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

- Note that independent variables can be time invariant.
- Individual (random) effects are independent:

$$\mathbb{E}(\mu_i, \mu_j) = 0 \qquad \text{if} i \neq j. \tag{14}$$

- Estimation method: GLS (*generalized least squares*).

- The error component (u_t):

$$u_{it} = \mu_i + \varepsilon_{it}, \tag{15}$$

where $\mu_i \sim \mathcal{N}(0, \sigma^2_\mu)$ and $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2_\varepsilon)$.

- Diagonal elements of the variance covariance matrix of the error term:

$$\begin{aligned} \mathbb{E}\left(u^2_{it}\right) &= \mathbb{E}\left(\mu^2_i\right) + \mathbb{E}\left(\varepsilon^2_{it}\right) + 2cov\left(\mu_i, \varepsilon_{it}\right) \\ &= \sigma^2_\mu + \sigma^2_\varepsilon \end{aligned}$$

- Non-diagonal elements of the variance covariance matrix of the error term ($t \neq s$):

$$cov\left(u_{it}, u_{is}\right) = \mathbb{E}\left(u_{it}u_{is}\right) = \mathbb{E}\left[\left(\mu_i + \varepsilon_{it}\right)\left(\mu_i + \varepsilon_{is}\right)\right].$$

- After manipulation:

$$cov\left(u_{it}, u_{is}\right) = \underbrace{\mathbb{E}\left(\mu^2_i\right)}_{\sigma^2_\mu} + \underbrace{\mathbb{E}\left(\mu_i\varepsilon_{it}\right)}_{0} + \underbrace{\mathbb{E}\left(\mu_i\varepsilon_{is}\right)}_{0} + \underbrace{\mathbb{E}\left(\varepsilon_{it}\varepsilon_{is}\right)}_{0} = \sigma^2_\mu$$

- Finally, variance covariance matrix of the error term for given individual (i):

$$\mathbb{E}\left(u_{i.}u_{i.}'\right)=\Sigma_{u,i}=\left(\sigma_{\mu}^2+\sigma_{\varepsilon}^2\right)\begin{bmatrix}1 & \rho & \dots & \rho \\ \rho & 1 & \vdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1\end{bmatrix}, \tag{16}$$

where

$$\rho=\frac{\sigma_{\mu}^2}{\sigma_{\mu}^2+\sigma_{\varepsilon}^2}.$$

- Note that Σ_u is block diagonal with equicorrelated diagonal elements $\Sigma_{u,i}$ but not spherical.
- Although disturbances from different (cross-sectional) units are independent presence of the time invariant random effects (μ_i) leads to equi-correlations among regression errors belonging to the same (cross-sectional) unit.

GLS estimator

- Consider the following linear model:

$$y = \beta X + u \tag{17}$$

when the variance covariance matrix of the error term ($\mathbb{E}(uu') = \sigma^2 I$) is spherical then the OLS estimator is motivated:

$$\beta^{OLS} = \left(X'X\right)^{-1} X'y. \tag{18}$$

- Let's relax the assumption about the variance covariance matrix of the error term, i.e., $\mathbb{E}(uu') = \Sigma$ and Σ is not diagonal but is positively defined. We can transform the baseline model (17) as follows:

$$Cy = C\beta X + Cu, \tag{19}$$

where $C = \Sigma^{-\frac{1}{2}}$.

- The variance covariance matrix of the error term becomes in the transformed model:

$$\mathbb{E}\left(CuC'u'\right) = \mathbb{E}\left(\Sigma^{-\frac{1}{2}} u(\Sigma^{-\frac{1}{2}})'u'\right) = \mathbb{E}\left(\Sigma^{-1}\Sigma\right) = I. \tag{20}$$

- After manipulations, the GLS estimator of the vector β is:

$$\beta^{GLS} = \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}y. \tag{21}$$

- **How to choose Σ?**.

- Using the GLS estimator we have:

$$\hat{\beta}^{RE} = \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}y, \quad (22)$$

$$Var\left(\hat{\beta}^{RE}\right) = \left(X'\Sigma^{-1}X\right)^{-1}, \quad (23)$$

but we don't know Σ !

- We know that $\Sigma = \mathbb{E}\left(uu'\right)$ is block-diagonal and:

$$cov\left(u_{it}, u_{js}\right) = \begin{cases} 0 & \text{if} \quad i \neq j, \\ \sigma_\mu^2 + \sigma_\varepsilon^2 & \text{if} \quad i = j \text{ and } s = t, \\ \rho\left(\sigma_\mu^2 + \sigma_\varepsilon^2\right) & \text{if} \quad i = j \text{ and } s \neq t, \end{cases}$$

where $\rho = \sigma_\mu^2 / \left(\sigma_\mu^2 + \sigma_\varepsilon^2\right)$.

- But we do not know σ_μ and σ_ε.
- There are many different strategies to estimates σ_μ and σ_ε.

Model estimation: RE – example

- We use the GLS transforms of the independent (x^*_{jit}) and dependent variable (y^*_{it}):

$$\begin{aligned} x^*_{jit} &= x_{jit} - \hat{\theta}_i \bar{x}_{ji} \\ y^*_{it} &= y_{it} - \hat{\theta}_i \bar{y}_i \end{aligned}$$

  where $\bar{x}_{ji}$ and $\bar{y}_i$ are the individuals means.
- Estimates of the transforming parameter:

$$\hat{\theta}_i = 1 - \sqrt{\frac{\hat{\sigma}^2_{\varepsilon}}{T_i \hat{\sigma}^2_{\mu} + \hat{\sigma}^2_{\varepsilon}}}.$$

- The estimates of the idiosyncratic error component σ_{ε}:

$$\hat{\sigma}^2_{\varepsilon} = \frac{\sum_i^n \sum_t^{T_i} \hat{u}^2_{it}}{NT - N - K + 1}$$

  where

$$\hat{u}^2_{it} = (y_{it} - \bar{y}_i + \bar{y}) - \hat{\alpha}^{Within} - (x_{it} - \bar{x}_i + \bar{x})\,\hat{\beta}^{Within},$$

  where $\hat{\alpha}^{Within}$ and $\hat{\beta}^{Within}$ stand for the within estimates.

- The error variance of individual specific random component (σ_μ^2):

$$\hat{\sigma}_\mu^2 = \frac{SSR^{Between}}{N-K} - \frac{\hat{\sigma}_\varepsilon^2}{\bar{T}}$$

where $\bar{T}$ is the harmonic mean of T_i, i.e., $\bar{T} = n/\sum_i^n (1/T_i)$, and $SSR^{Between}$ stands for the sum of squared residuals from the between regression (details on further lectures):

$$SSR^{Between} = \sum_i^n \left(\bar{y}_i - \hat{\alpha}^{Between} - \bar{x}_i \hat{\beta}^{Between}\right)$$

where $\hat{\beta}^{Between}$ and $\hat{\alpha}^{Between}$ stand for the coefficient estimates from the between regression.

Model estimation: RE – Swamy-Arora method

- Method which gives more precise estimates in small samples and unbalanced panels.
- The estimation of σ_ε (the idiosyncratic error component) is the same
  $\Longrightarrow$ it bases on the residuals from the within regression.
- The variance or the individual error term:

$$\hat{\sigma}^2_{\mu,SA} = \frac{SSR^{Between} - (N-K)\,\hat{\sigma}^2_\varepsilon}{NT - tr},$$

where $SSR^{Between}$ is the sum of the squared residuals from the between regression and

$$tr = \text{trace}\left\{\left(X'PX\right)^{-1}X'ZZ'X\right\}$$

$$P = \text{diag}\left\{\frac{1}{T_i}e_{T_i}e'_{T_i}\right\}$$

$$Z = \text{diag}\left\{e_{T_i}\right\}$$

where e_{T_i} is a $T_i \times 1$ vector of ones.

Testing on the random effect

- In the standard RE model, the variance of the random effect is assumed to be σ_μ^2 ($\mu \sim \mathcal{N}\left(0, \sigma_\mu^2\right)$).
- We can test for the presence of heterogeneity:

$$\begin{array}{lcl} \mathcal{H}_0: & \sigma_\mu & = & 0 \\ \mathcal{H}_1: & \sigma_\mu & \neq & 0 \end{array}$$

- If the null hypothesis is rejected, then we conclude that there are random individual differences among sample members, and that the random effects model is appropriate.
- If we fail to reject the null hypothesis, then we have no evidence to conclude that random effects are present.
- We construct the Lagrange multiplier statistic:

$$LM = \sqrt{\frac{NT}{2(T-1)}} \left(\frac{\sum_{i=1}^{N} \left(\sum_{t=1}^{T} \hat{u}_{it} \right)^2}{\sum_{i=1}^{N} \sum_{t=1}^{T} \hat{u}_{it}^2} - 1 \right), \tag{24}$$

  where $\hat{u}_{i,t}$ stands for the residuals, i.e., $\hat{u}_{it} = y_{it} - \hat{\alpha}_0 - \hat{\beta}_1 x_{1it} - \ldots - \hat{\beta}_k x_{kit}$.
- Conventionally, $LM \sim \chi^2(1)$. In large samples, $LM \sim \mathcal{N}(0, 1)$.