Panel Econometrics

Paulo M. M. Rodrigues

February 2022

Heterogeneous panel data models

Outline

Heterogeneity in the slope coefficients I

- In the standard linear panel data model we control for unobserved heterogeneity:

$$y = \beta' X + u \quad (1)$$

  where u is the sum of individual-specific component (in the RE model) and the idiosyncratic component.
- In the FE model, the individual-specific intercepts are introduced while the u contains only the idiosyncratic shock.
- At the same time, we have assumed that all slope coefficients (vector β) are the same for all unit and all periods.
- In the above formulation, we do not allow for any interaction between individual effects and explanatory variable.

Heterogeneity in the slope coefficients I

- Consider the following formulation:

$$y_{it} = \beta_{it}' X_{it} + u_{it} \tag{2}$$

where all slope coefficients captured by β_{it} are now time-varying and individual-specific.

  - Although the above general formulation seems to be more realistic it lacks any explanatory power and is not useful for prediction.
  - The above model is not estimable since the number of parameters exceeds the number of observations.

More applicable formulations:

$$\begin{aligned} y_i &= \beta_i' X_i + u_i, \\ y_t &= \beta_t' X_t + u_t \end{aligned}$$

Heterogeneity in the slope coefficients II

- Which kind of heterogeneity in the slopes should we introduce?
- In general, we pay more attention to individual effects but it depends on
  - T and N,
  - the research question.
- To account for the individuals differences in the slope coefficients we introduce three approaches:
  - Seemingly Unrelated Regression (SUR),
  - Swamy's random coefficient model,
  - Mean group estimation.

Seemingly Unrelated Regression (SUR)

Seemingly Unrelated Regression (SUR) I

- **Seemingly Unrelated Regression (SUR)** is an estimation method that is designed to estimate a system of linear equations (with potentially a different set of explanatory variables) and which accounts for the cross-equation correlation of the error term.
- Consider the following set of equations:

$$y_i = X_i\beta_i + \varepsilon_i, \text{ for } i = 1, ..., m$$

where the index i denotes the i-th equation in the considered system.
- In the matrix form:

$$\begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_m \end{bmatrix} = \begin{bmatrix} X_1 & 0 & ... & 0 \\ 0 & X_2 & ... & 0 \\ ... & ... & ... & \\ 0 & 0 & ... & X_m \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ ... \\ \beta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ ... \\ \varepsilon_m \end{bmatrix}$$

- In the i-th equation, K_i parameters are estimated. It yields the total number of coefficient $K = \sum_{i=1}^{m} K_i$. In addition, $K_i \geq T_i$
- Strictly exogeneity is assumed, i.e., $\mathbb{E}(\varepsilon | X_1, ..., X_m) = 0$.

Seemingly Unrelated Regression (SUR) I

- In the SUR framework, it is possible to assume that the covariance matrix of the error term is not diagonal:

$$\Omega = \mathbb{E}(\varepsilon\varepsilon'|X_1, ..., X_m) = \begin{bmatrix} \sigma_{11}^2\mathbf{I} & \sigma_{12}^2\mathbf{I} & ... & \sigma_{1m}^2\mathbf{I} \\ \sigma_{21}^2\mathbf{I} & \sigma_{22}^2\mathbf{I} & ... & \sigma_{2m}^2\mathbf{I} \\ ... & ... & ... & \\ \sigma_{m1}^2\mathbf{I} & \sigma_{m2}^2\mathbf{I} & ... & \sigma_{mm}^2\mathbf{I} \end{bmatrix}$$

Seemingly Unrelated Regression (SUR) II

- Given the above structure of the variance-covariance matrix of the error term, the system of equations can be estimated with FGLS (feasible generalized least squares). Conventionally, the two-step estimation includes the following steps:
  1. Run the OLS regression for the considered system of equations to get consistent and unbiased estimates of the variance-covariance matrix of the error term ($\widehat{\Omega}$).
  2. Based on the estimates of the $\widehat{\Omega}$, standard GLS estimation can be applied:

$$\widehat{\beta}^{SUR} = (X'\widehat{\Omega}^{-1}X)^{-1}X'\widehat{\Omega}^{-1}y$$

- Note that if Ω is diagonal then $\widehat{\beta}^{SUR}$ will be close to the OLS estimator.

The SUR estimation and panel data

- In the context of long and narrow panel data, the SUR can be applied to account for a potential heterogeneity in the slopes.
- Consider the case of long (relatively large T) and narrow (not so large N) panel. Then, the standard linear model can be expressed as a set of equations:

$$\begin{aligned} y_1 &= \beta_1' X_1 + \varepsilon_1, \\ y_2 &= \beta_2' X_2 + \varepsilon_2, \\ ... &= ... \\ y_N &= \beta_N' X_N + \varepsilon_N, \end{aligned} \tag{0.1}$$

  where β_i is the individual-specific vector of the structural parameters.
- The SUR method accounts for cross-equation correlation. In the above case, this correlation is equivalent to cross-sectional dependence.

The SUR estimation and panel data

- It is possible to test for slope heterogeneity. The standard Wald test can be used to verify the hypothesis about:
  - homogeneity of all slopes, i.e., $H_0 : \beta_1 = ... = \beta_N$, where β_i stands for the vector of parameters of the i^{th} unit.
  - homogeneity of some slopes, i.e., $H_0 : \beta_{1,j} = ... = \beta_{N,j}$, where $\beta_{i,j}$ stands for j^{th} parameter for $i - th$ unit.

Testing cross-equation (cross-sectional) correlation of the error term

- **The SUR method** provides more efficient estimates since it accounts for cross- equation dependence.
- Cross-equation dependence can be tested with the LM statistic (Breusch and Pagan, 1980):

$$LM = T \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \widehat{\rho}_{i,j}^{2}, \tag{0.2}$$

where $\rho_{i,j}^2$ is cross-sectional correlation coefficient:

$$\widehat{\rho}_{i,j}, = \frac{\sum_{t=1}^{T} \widehat{\varepsilon}_{it}\widehat{\varepsilon}_{jt}}{(\sum_{t=1}^{T} \widehat{\varepsilon}_{it})^{1/2}(\sum_{t=1}^{T} \widehat{\varepsilon}_{jt})^{1/2}} \tag{0.3}$$

The LM statistic is valid for fixed N as $T \to \infty$ and is asymptomatically distributed as a χ^2 with $N(N-1)/2$ degrees of freedom.

Swamy's random coefficient model

Swamy's random coefficient model I

- Swamy (1970) proposes the random coefficient model.
- Consider the following model:

$$y_i = X_i \beta_i + \varepsilon_i$$

where the individual-specific slope β_i is the sum of common (β) and unit-specific (α_i) components, i.e.,

$$\beta_i = \beta + \alpha_i,$$

where

$$\begin{aligned} \mathbb{E}(\alpha_i) &= 0 \\ \mathbb{E}(\alpha_i \alpha_i') &= \Sigma \end{aligned}$$

Swamy's random coefficient model I

- **Question:** How can β and Σ be estimated?
- The dependent variable can be expressed:

$$y_i = X_i\beta_i + \varepsilon_i = X_i\beta + X_i\alpha_i + \varepsilon_i = X_i\beta + \nu_i,$$

where $\nu_i = X_i\alpha_i + \varepsilon_i$ and $(E)(\nu_i) = 0$.

- The variance-covariance of the error term ν_i for the i^{th} unit is the following:

$$\mathbb{E}(\nu_i\nu_i') = \mathbb{E}((X_i\alpha_i + \varepsilon_i)(X_i\alpha_i + \varepsilon_i)') = \mathbb{E}(\varepsilon_i\varepsilon_i') + X_i\mathbb{E}(\alpha_i\alpha_i')X_i'.$$

- If the idiosyncratic error term is spherical (homoscedastic) then:

$$\mathbb{E}(\nu_i\nu_i') = \sigma_i^2\mathbb{I} + X_i\Sigma X_i' = \Pi_i$$

Swamy's random coefficient model II

- The Π variance-covariance matrix for the error term will be block-diagonal.
- Finally, the GLS estimator can be applied:

$$\widehat{\beta}^{RC} = \left(\sum_i X_i' \Pi_i^{-1} X_i \right)^{-1} \sum_i X_i' \Pi_i^{-1} y_i = \sum_i W_i \widehat{\beta}_i^{OLS}$$

where $\widehat{\beta}_i^{OLS}$ is the unit-specific OLS estimates and W_i:

$$W_i = \left[\sum_i (\Sigma + V_i)^{-1} \right]^{-1} (\Sigma + V_i)^{-1}$$

where V_i is the panel-specific variance-covariance of $\widehat{\beta}_i^{OLS}$, i.e., $\widehat{V} = \sigma_i^2 (X_i' X_i)^{-1}$.

- The variance of β can be calculated as:

$$Var(\beta) = \sum_i (\Sigma + V_i)^{-1}$$

Swamy's random coefficient model III

- Finally, the remainder element of the variance-covariance components which captures the variation of the slope coefficients, i.e., Σ, can be estimated based on the variation in the panel-specific β^{OLS} estimates:

$$\widehat{\Sigma} = \frac{1}{N-1}\left(\Sigma_i \widehat{\beta}_i^{OLS}\left(\widehat{\beta}_i^{OLS}\right)' - N\bar{\beta}^{OLS}\left(\bar{\beta}^{OLS}\right)'\right) - \frac{1}{N}\sum_i \widehat{V}_i$$

  where $\bar{\beta}^{OLS}$ is the average of the OLS estimates.

- Swamy (1970) postulates to omit the last component because it is negligible in large samples and it can be not positive definite.

Testing homogeneity in slopes

- To test whether the random coefficient model is statistically motivated one might compare the panel-specific estimates with their weighted (by V^{-1}) average.
- Test statistic:

$$\mathcal{T} = \sum_{i=1}^{N} \left(\widehat{\beta}_i^{OLS} - \tilde{\beta}\right)' \widehat{V}_i^{-1} \left(\widehat{\beta}_i^{OLS} - \tilde{\beta}\right)$$

where

$$\tilde{\beta} = \left(\sum_{i=1}^{N} \widehat{V}_i^{-1}\right)^{-1} \sum_{i=1}^{N} \widehat{V}_i^{-1})^{-1} \widehat{\beta}_i^{OLS}$$

- The null hypothesis:

$$H_0 : \beta_1 = \beta_2 = ... = \beta_N$$

.

- The test statistic $\mathcal{T}$ is asymptotically χ^2 distributed with $k \times (N-1)$ degrees of freedom.

Mean group estimation

Mean group estimation I

- The Mean Group estimator (MG) was proposed by Pesaran and Smith (1995) to deal with dynamic random coefficient models.
- The MG estimator is defined as the average of the unit-specific OLS estimators $\widehat{\beta}_i^{OLS}$:

$$\widehat{\beta}^{MG} = \frac{1}{N} \sum_{i=1}^{N} \widehat{\beta}_i^{OLS}$$

  where

$$\widehat{\beta}_i^{OLS} = (X_i' X_i)^{-1} X_i' y_i$$

- It is assumed that all explanatory variables are strictly exogenous.
- The MG estimation is used when both T and N are sufficiently large.

Mean group estimation II

- The MG estimation can be applied irrespectively of the nature of heterogeneity in the slope coefficient. It can be applied if
    1. the differences in slopes are random (as in the Swamy estimator),
    2. diversity in the slopes can be captured by the fixed effects.
- The variance of the MG estimator:

$$Var(\widehat{\beta}^{MG}) = \frac{1}{N(N-1)} \sum_{i=1}^{N} (\widehat{\beta}_i^{OLS} - \widehat{\beta}^{MG})(\widehat{\beta}_i^{OLS} - \widehat{\beta}^{MG})'.$$

- The MG estimator will be very close to the Swamy's estimator if T tends to infinity and there is some heterogeneity in the slopes:

$$\lim_{T \to \infty} (\widehat{\beta}^{MG} - \widehat{\beta}^{RC}) = 0.$$

Pooled MG

- In the pooled mean group estimation all coefficients are pooled, i.e, they are constrained to be identical:

$$\forall_i \beta_i = \beta$$

- However, one might pool only a subset of coefficients.
- To test the assumption about homogeneity of coefficients one can use the standard Hausman test comparing mean group and pooled mean group estimates:
  - Under null both estimates are consistent while under alternative only the mean group estimates are consistent.
  - By pooling we increase the efficiency of the estimates.

Cross-sectional dependence and CCE estimator

- Cross-sectional dependence leads to endogeneity in the mean group estimation. Recalling the least square estimator:

$$\widehat{\beta}_i^{LS} = \beta_i + (X'X)^{-1}X'\mathbb{E}(\varepsilon_i),$$

  one may observe that the presence of common factors lead to endogeneity.
- In the CCE (common correlated effects) estimation we control for the multi-factor structure of the error term. To account for common factors, individual-specific regressions are extended by cross-sectional averages of the dependent variable.
- Given a high degree of uncertainty about the structure of the error term one can also use cross-sectional averages of the explanatory variable. In addition, the lags of cross-sectional averages of both the dependent and explanatory variables can be also included (see Chudik and Pesaran, 2015).