Panel Data

w. Vicente Royuela Mora

Universitat de Barcelona

Complementary course notes

Thor Donsby Noe *

November 14, 2018

*Department of Economics, University of Copenhagen, Øster Farimagsgade 5, DK-1353 Copenhagen K, Denmark (e-mail: jwz766@alumni.ku.dk)

Contents

"..."

.

1 Introduction

Example macro panel data:
Maddison, 2007, and IFS

- Formatting data
- Describtive statistics
- Graphs
- Maps

2 Fixed Effects estimators

2.1 Simultaneous equations models with exogenous explanatory variables

Four different models:

M1: $\alpha \neq, \beta \neq$. Estimation individual by individual (GLS).

M2: $\alpha =, \beta =$. Equal constant terms and slopes (presence of homogeneity).

M3: $\alpha \neq, \beta =$. Equal slopes, different constant terms.

M4: $\alpha =, \beta \neq$. Equal constant terms, different slopes.

Choosing between them using
Test for Homogeneity:

a. Estimate the extended/unrestricted model.

b. Estimate the restricted model.

c. H_0 : Homogeneity (the unrestricted is not better than the restricted).
Reject H_0 if the F-value is higher than the critical value of the F-distribution.

$$F = \frac{(SSR_R - SSR_{UR})/r}{SSR_{UR}/df}$$

In ML we maximize the probability.
In OLS we don't care about the variance.

2.2 The fixed effects model

α_i is a parameter capturing the individual effect (time-invariant!).

$$y_{it} = i\alpha_i + X_{it}\beta + \varepsilon_{it}$$
$$\Rightarrow y = [\ i \quad X\ ] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \varepsilon$$

Leasy Squares Dummy Variables (LSDV):

- FE model: All variables are within-transformed e.g the deviations from the mean.

Test the homogeneity analysis:

$$H_0 : \ \alpha_1 = \alpha_2 = \cdots = \alpha_M$$

2.3 Within and between estimators

The overall variance: Weighted variation between the within-variance and between-variance.

2.4 Effects of group and time

3 Random effects estimator

3.1 The random effects model

3.2 The generalized least squares estimation

3.3 Feasible Generalized Lest squares (unkown)

4

4 Fixed effects vs. random estimator

	No endogeneity	Endogeneity
R.E.	Consist. & Efficient	Inconsistent
F.E.	Consist. but inefficient	Consist. but ineff.

Table 1: Endogeneity problems

4.1 The Breush Pagan test

4.2 The Hausman test

- FE - β^{FE}
- FE - β^{RE}
- $()\beta^{FE} - \beta^{RE})$?

Problems:

- Hausman assumptions
- We cannot use "fixed" variables (no time variation).

The Mundlak estimation:

→ Allows joint estimation F.E. / BE

- Diff(FE-BE) → estimate 'pseudo' Hausman test.

→ Allows for including 'fixed' variables (with no time-variation)

RE:

$$y_{it} + x_{it}\beta + \psi_i + \varepsilon_{it}$$

Mundlak regression:

Including both RE and BE.

$$y_{it} + (x_{it} - \bar{x}_i)\beta^w + \bar{x}_i\beta^b + \psi_i + \varepsilon_{it}$$

Is just an instrumental regression!

Standard errors are unreliable and R^2 is lower.

- FE: We capture anything permanent for any individual in α_i
  - → Get rid of anything permanent.
  - → Good for controlling.
  - – We don't know what the individual effects mean though!

- Mundlak: We 'only' capture anything permanent considered in $\bar{x}_i$

Increasing the data set (information)

- Hausman test: Is likely to find endogeneity though the difference in coefficients is very small.

4.3 Long run vs. short run effects

Baltagi & Griffin (1984) "Short and Long Run Effects in Pooled Models".

- Betweeen est. $\rightarrow$ LR effect.
- Within est. $\rightarrow$ SR effect.
- OLS and RE $\rightarrow$ average of SR & LR effect.

Inequality and growth

- Positive in the SR
- Negative in the LR

5 Heteroskedasticity and autocorrelation in panel data

5.1 Heteroskedasticity in FE model

$$y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it} \tag{5.1}$$

Heteroscedasticity over time: σ^2_ε

5.2 Heteroskedasticity in RE model

$$y_{it} = x_{it}\beta + \xi_i + \varepsilon_{it} \tag{5.2}$$

Heteroscedasticity over individuals: σ^2_ξ
Heteroscedasticity over time: σ^2_ε

5.3 Autocorrelation in the FE model

In eq. 5.1 it can be that

$$\varepsilon_{it} = V_{it} + \theta V_{it-1}$$
$$\rightarrow \text{AR(1)} \rightarrow \varepsilon_{it} = \rho_1 \varepsilon_{it-1} + v_{it}$$

Assumption: Everyone is homogenous in their autocorrelation parameter.

Consistency can come from either a high number of N or T.

5.4 Autocorrelation in the RE model

We need to start with the FE estimation

$\rightarrow$ To get a consistent estimate of ρ.

- Proceed to estimate either a FE or RE model.

6 Incomplete/unbalanced panels

Incomplete panels are similar to heteroscedasticity issues.

- Some issues can be fixed with weighting observations.
- For some procedures we will need complete panels though.

7 Dynamic panels

Autocorrelation in y_{it}.

$$y_{it} = \alpha y_{it-1} + \alpha_{it}\beta + \cdots$$

GMM estimation.