Vrije Universiteit Amsterdam

Econometrics for Quantitative Risk Management

Panel Data

Assignment III

Authors
Joris Kreulen (2653365) & Aytek Mutlu (2648232)

October 14, 2019

1 Introduction

In this assignment we will work with panel data. We will explore various methods of working with this type of data, i.e. the implementation of a 'Fixed Effects model', 'Pooled OLS' and a 'Random Effects model'. The output of all models shall be presented and the differences amongst models discussed. The dataset we use is from Hansen (2019). The set contains various variables about company characteristics in different industries.

2 Fixed Effects Model

To estimate our fixed effects model we first define our model:

$$y_{it} = X_{it}\beta + \alpha_i + \upsilon_{it}$$

Where y_{it} is the dependent variable observed for individual i at time t. X_{it} is the time-variant $1/k$ (*the number of independent variables*) regressor vector. β is the $k \times 1$ matrix of parameters. α_i is the unobserved time-invariant individual effect, in this case company specific characteristics. υ_{it} is the error term.

The fixed effects (FE) model allows α_i to be correlated with the regressor matrix X_{it}. Strict exogeneity with respect to the idiosyncratic error term υ_{it}, however, is still required. Since α_i is not observable, it cannot be directly controlled for. Our FE model will eliminate α_i by demeaning the variables as follows:

$$y_{it} = (X_{it} - \overline{X}_{it})\beta + (\alpha_i - \overline{\alpha}_i) + (\upsilon_{it} - \overline{\upsilon}_{it})$$

Where the term $(\alpha_i - \overline{\alpha}_i)$ disappears since α_i is a constant. This way we can estimate the vector β without the need to estimate $n-1$ dummies and an intercept. In addition, this basically gets rid of all between-subject variability (which may be contaminated by omitted variable bias) and leaves only the within-subject variability to analyze. It should be noted that demeaning is performed per company.

The estimated $\beta's$ and standard errors are found in the table below:

	Beta	**Standard Error**
vala	0.0017	0.000760
debta	-0.0139	0.003692
cfa	0.0491	0.008933

We compare these with Table 17.2 in Hansen (*2019, august*):

As it can be observed, our beta estimations are exactly the same where standard errors are slightly underestimated even with the adjustment of degrees of freedom.

	Beta		Standard Error	
	Our Results	**Hansen**	**Our Results**	**Hansen**
vala	0.0017	0.0017	0.000760	0.0008
debta	-0.0139	-0.0139	0.003692	0.0049
cfa	0.0491	0.0491	0.008933	0.0132

3 Pooled OLS

This is the most 'basic' model for analyzing panel data. The method implies estimating OLS without recognizing the panel data structure, thus ignoring individually specific effects, in our case this means we don't recognize the individual companies and the time element. The set up is more of a cross-sectional regression type. This would however mean that the intercept value is the same across units or entities (in this case, companies) and that the slope coefficients in the vector β are constant across the companies. These assumptions of a constant intercept and slope coefficients are highly restricted and far-fetched, and so we expect high standard deviations for the β estimates in our regression output.

Since we're now working with classic OLS, assumptions must be made:

- The explanatory variables are exogenous, ensuring OLS will be unbiased and consistent
- The error terms are homoscedastic and not autocorrelated, ensuring OLS will be efficient.

The model to be run is as follows:

$$y = \alpha + X\beta + D\alpha + \eta$$

where α is the constant mean, D are the industry dummies and η is the residual.

Here are the results and comparison with Hansen:

	Beta		Standard Error	
	Our Results	**Hansen**	**Our Results**	**Hansen**
vala	0.0024	0.0024	0.001	0.001
debta	0.0096	0.0096	0.003	0.0041
cfa	0.0261	0.0261	0.007	0.0111
nyseamex	-0.0167	-0.0167	0.001	0.0024

4 Random Effects Model

For the Fixed Effects model, we dropped the assumption $X_i \perp \alpha_i$, however, if $X_i \perp \alpha_i$ is the case then the individual effect α_i can be considered a *random* *effect* w.r.t. regressors X_i. We must implement a Random Effects model when this holds.

If this random effects assumption holds, the RE model is more efficient than the FE model. However, if this assumption does not hold, the RE model is not consistent. The Durbin–Wu–Hausman test is often used to discriminate between the fixed and the random effects model *[1]*.

We have followed the cookbook in the slides with the following steps:

1. Matrices are filled with maximum year data to handle unbalanced panel among groups,

2. OLS is applied to find pooled estimators,

3. Covariance matrix of each group calculated from residuals of pooled estimator is averaged,

4. Random Effect estimators are calculated from this averaged covariance matrix.

   Unfortunately, this procedure did not end up with similar results with Hansen. We will investigate the problem try to find out the missing, wrong or extra steps. Here are the estimators:

	Our Results	**Hansen**
vala	0.0029	0.0019
debta	0.0899	-0.0092
cfa	0.0547	0.0412
nyseamex	0.0233	-0.0181

References

[1] Nerlove, Marc *Essays in Panel Data Econometrics. (pp. 36–39)*. Cambridge University Press., 2005.