AFRE 835: Introductory Econometrics

Chapter 14: Advanced Panel Data Methods

Spring 2017

Introduction

- In the last chapter, we saw that panel data provided a mechanism for controlling for time invariant unobserved heterogeneity.
- This chapter introduces two additional approaches:
  1. The **fixed effects (FE)** model or transformation.
  2. The **random effects (RE)** model.
- We will not cover the section on the **Correlated Random Effects** approach, which is in some sense a compromise between RE and FE models.

Outline

1. Fixed Effects Estimation

2. Random Effects Models

The Fixed Effects Transformation

- First differencing provided one way of eliminating the unobserved heterogeneity term a_i.
- An alternative approach is to **time-demean** the data.
- Consider the following simple regression model for a panel data set

$$y_{it} = \beta_0 + \beta_1 x_{it} + a_i + u_{it} \quad t = 1, \ldots, T. \tag{1}$$

- For each i, we can compute the mean value of this equation over time; i.e.,

$$\bar{y}_i = \beta_0 + \beta_1 \bar{x}_i + a_i + \bar{u}_i. \tag{2}$$

where $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$, $\bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$, and $\bar{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_{it}$.

- Subtracting (2) from (1) yields

$$y_{it} - \bar{y}_i = \beta_1(x_{it} - \bar{x}_i) + u_{it} - \bar{u}_i$$

$$\text{or} \qquad \ddot{y}_{it} = \beta_1 \ddot{x}_{it} + \ddot{u}_{it} \tag{3}$$

where $\ddot{y}_{it} = y_{it} - \bar{y}_i$ denotes the time-demeaned value of y_{it}, etc.

The Fixed Effects Transformation (cont'd)

- If we apply the pooled OLS estimator to the model in (3), the OLS estimator β_1 will be unbiased if

  $$E(\ddot{u}_{it}|\ddot{x}_{it}) = 0 \tag{4}$$

  which is effectively going to require strict exogeneity of our original regressor; i.e.,

  $$E(u_{it}|\boldsymbol{x}_i) = 0 \tag{5}$$

- The pooled OLS estimator applied to the time demeaned data is referred to as the **fixed effects (FE) estimator**.
- Notes
  - We will also need x_{it} to vary over time for at least one individual or else $\ddot{x}_{it} = 0 \;\; \forall i, t$ and β_1 will not be identified.
  - There is no intercept in the demeaned model in (3).
  - The model places *no* restrictions on the relationship between x_{it} and a_i.

Multiple Regressors

- Extending the FE estimator is straightforward.
- Given a linear regression model

  $$y_{it} = \beta_0 + \beta_1 x_{it1} + \cdots + \beta_k x_{itk} + a_i + u_{it} \quad t = 1, \ldots, T. \tag{6}$$

  ...the FE estimator corresponds to applying pooled OLS to the demeaned equation:

  $$\ddot{y}_{it} = \beta_1 \ddot{x}_{it1} + \cdots + \beta_1 \ddot{x}_{itk} + \ddot{u}_{it} \tag{7}$$

- As in the simple regression model, each regressor has to vary over time for at least one individual.
- Note: The degrees of freedom for the FE estimator is $NT - N - k$, because of the k slope parameters estimated and the time demeaning for N individuals.

  ...Another way to think of it is that for each individual, one of their demeaned observations is redundant.

Formal Conditions for Unbiasedness and Consistency

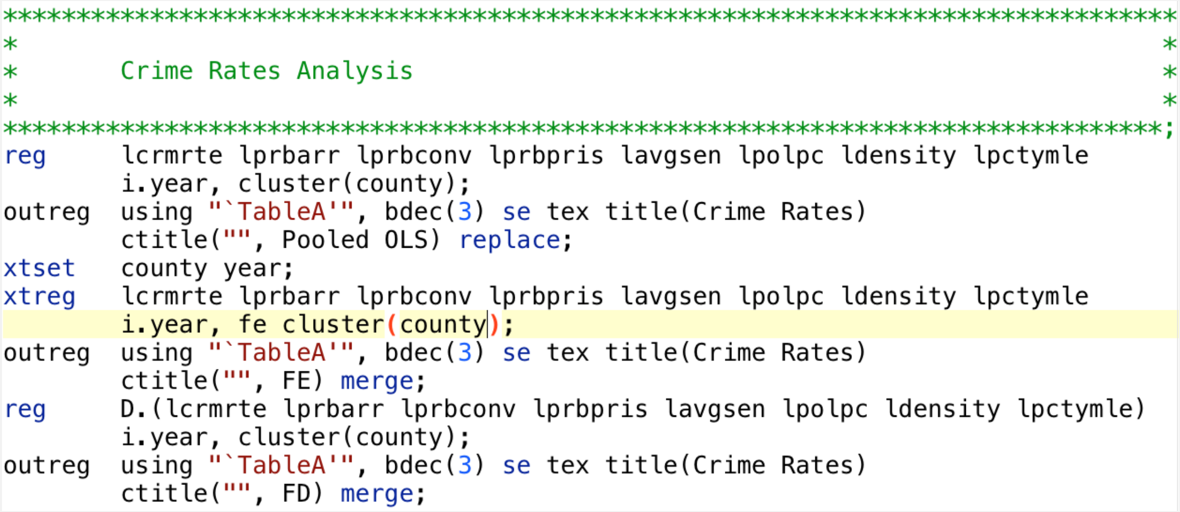
- Assumptions required for consistency
  1. The model is linear in the parameters and a_i (as in (6));
  2. We have a random sample in the cross-section;
  3. Each independent variable varies over time for at least one individual i;
  4. Strict exogeneity holds; i.e.,

$$E(u_{it}|x_{i1},\dots,x_{iT},a_i)=0 \tag{8}$$

- These same assumptions are also sufficient for the FE estimator to be consistent with a fixed T as $N \to \infty$.

Example #1 (Code): Crime Rates in NC

- Data set: Crime4.dta
- Here is the code for pooled OLS, FE, and FD estimation with time dummies



```
*********************************************************************************
*                                                                               *
*       Crime Rates Analysis                                                    *
*                                                                               *
********************************************************************************;
reg     lcrmrte lprbarr lprbconv lprbpris lavgsen lpolpc ldensity lpctymle
        i.year, cluster(county);
outreg  using "`TableA'", bdec(3) se tex title(Crime Rates)
        ctitle("", Pooled OLS) replace;
xtset   county year;
xtreg   lcrmrte lprbarr lprbconv lprbpris lavgsen lpolpc ldensity lpctymle
        i.year, fe cluster(county);
outreg  using "`TableA'", bdec(3) se tex title(Crime Rates)
        ctitle("", FE) merge;
reg     D.(lcrmrte lprbarr lprbconv lprbpris lavgsen lpolpc ldensity lpctymle)
        i.year, cluster(county);
outreg  using "`TableA'", bdec(3) se tex title(Crime Rates)
        ctitle("", FD) merge;
```

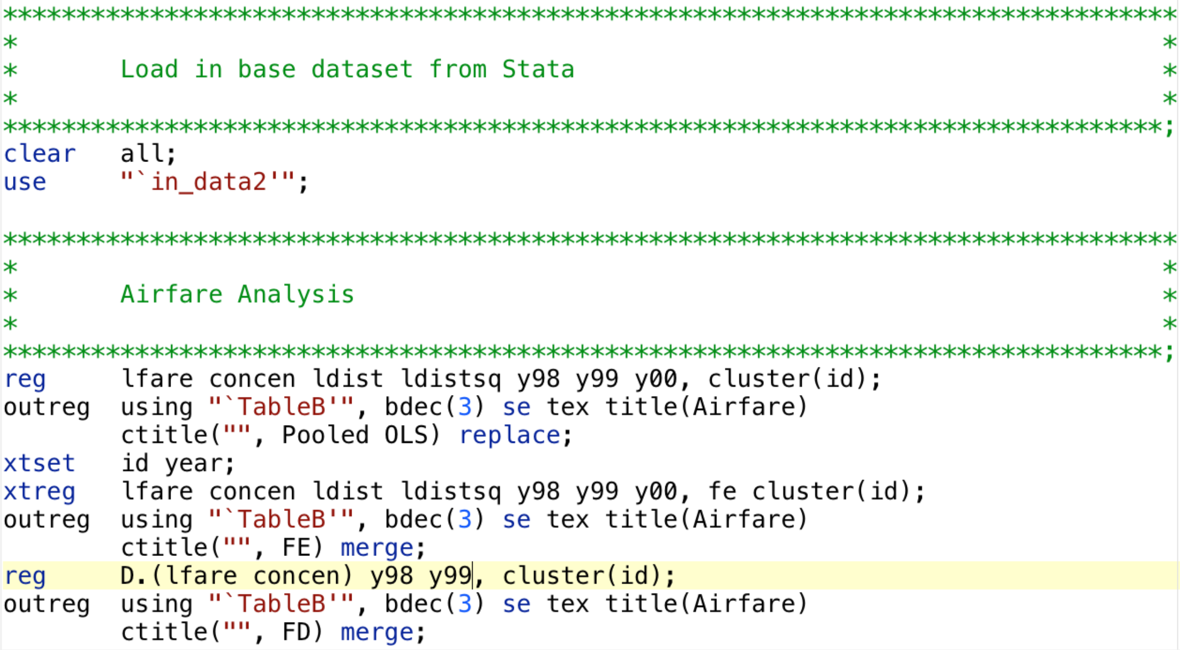
Ex. #1 (Results): Crime Rates in NC (with time dummies)

	Pooled OLS	FE	FD
lprbarr	-0.512 (0.137)**	-0.358 (0.059)**	-0.327 (0.056)**
lprbconv	-0.389 (0.085)**	-0.283 (0.050)**	-0.238 (0.040)**
lprbpris	0.110 (0.083)	-0.181 (0.045)**	-0.165 (0.046)**
lavgsen	-0.121 (0.105)	-0.004 (0.033)	-0.022 (0.026)
lpolpc	0.283 (0.139)*	0.419 (0.086)**	0.397 (0.104)**
ldensity	0.250 (0.067)**	0.372 (0.396)	0.134 (0.417)
lpctymle	0.053 (0.132)	0.359 (0.537)	0.932 (0.630)
_cons	-2.139 (0.986)*	-0.756 (1.248)	0.020 (0.014)
R^2	[illegible]	[illegible]	[illegible]
N	630	630	540

* $p < 0.05$; ** $p < 0.01$

Example #2 (Code): Impact of Concentration on Air Fare

- Data set: airfare.dta
- Code for pooled OLS, FE, and FD estimation with time dummies



```
*****************************************************************************
*                                                                           *
*       Load in base dataset from Stata                                     *
*                                                                           *
****************************************************************************;
clear   all;
use     "`in_data2'";

*****************************************************************************
*                                                                           *
*       Airfare Analysis                                                    *
*                                                                           *
****************************************************************************;
reg     lfare concen ldist ldistsq y98 y99 y00, cluster(id);
outreg  using "`TableB'", bdec(3) se tex title(Airfare)
        ctitle("", Pooled OLS) replace;
xtset   id year;
xtreg   lfare concen ldist ldistsq y98 y99 y00, fe cluster(id);
outreg  using "`TableB'", bdec(3) se tex title(Airfare)
        ctitle("", FE) merge;
reg     D.(lfare concen) y98 y99, cluster(id);
outreg  using "`TableB'", bdec(3) se tex title(Airfare)
        ctitle("", FD) merge;
```

Ex. #2 (Results): Airfare w/time dummies

	Airfare		
	Pooled OLS	FE	FD
concen	0.360 (0.059)**	0.169 (0.049)**	0.176 (0.043)**
ldist	-0.902 (0.272)**		
ldistsq	0.103 (0.020)**		
y98	0.021 (0.004)**	0.023 (0.004)**	-0.039 (0.006)**
y99	0.038 (0.005)**	0.036 (0.005)**	-0.048 (0.005)**
y00	0.100 (0.006)**	0.098 (0.006)**	
_cons	6.209 (0.912)**	4.953 (0.030)**	0.061 (0.003)**
R^2	0.41	0.14	0.04
N	4,596	4,596	3,447

The Dummy Variable Regression

- While the FE estimator can be obtained using the two-step process; i.e.,
  1. Demeaning the dependent and independent variables,
  2. Regressing $\ddot{y}_{it}$ on the $\ddot{x}_{itj}$'s.

  . . . One can also obtain the *numerically identical* estimates using what is known as the **dummy variable regression**, treating each a_i as a parameter to estimate via pooled OLS where

$$y_{it} = \sum_{j=1}^{N} a_j \delta_{ij} + \beta_1 x_{it1} + \cdots + \beta_k x_{itk} + u_{it} \quad t = 1, \ldots, T. \qquad (9)$$

  where $\delta_{ij} = 1$ if $i = j$, $=0$ otherwise.

- It is important to know, however, that the OLS estimator for the individual fixed effects $\hat{a}_i$ will not be consistent, since T is treated as fixed and is typically small.

Fixed Effects vs. First Differencing

- When $T = 2$, FE and FD estimators are identical.
- This is no longer true for $T > 2$.
- In general, FE will be more efficient when the u_{it} are not correlated, whereas FD will dominate when there greater serial correlation.
- Both estimators can be sensitive the measurement errors.
- It makes sense in most settings to estimate both models to see if the results are sensitive to this choice.

Random Effects Models

- The advantage of the FE estimator is that it purges the model of the potential bias and inconsistencies induced by the unobservable heterogeneity a_i being correlated with the regressors of the model.
- Such correlation would violate the zero conditional mean assumption $E(v_{it}|\boldsymbol{x}_i) = 0$.
- However, *if* indeed the a_i are not correlated with the independent variables of our model, then the FE estimator will be inefficient.

  ... This inefficiency comes from removing some of the variation in the model by demeaning the data prior to applying pooled OLS.
- The *random effects (RE) model* avoids this inefficiency by using the additional assumption $Cov(x_{itj}, a_i) = 0 \ \ \forall i = 1, \ldots, N; t = 1, \ldots, T; j = 1, \ldots, k.$
- Appendix 14A provides the full set of assumptions required for consistency and asymptotic normality of the RE estimator.

Random Effects Models (cont'd)

- The random effects model starts with the specification

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \cdots + \beta_k x_{itk} + v_{it} \quad t = 1, \ldots, T. \tag{10}$$

where $v_{it} = a_i + u_{it}$ denotes the **composite error term**, where a_i and u_{it} are uncorrelated, with $Var(a_i|\boldsymbol{x}_i) = \sigma_a^2$ and $Var(u_{it}|\boldsymbol{x}_i) = \sigma_u^2$.

- These assumptions imply that $Var(v_{it}|\boldsymbol{x}_i) = \sigma_a^2 + \sigma_u^2$ and $Corr(v_{it}, v_{is}) = \sigma_a^2/(\sigma_a^2 + \sigma_u^2)$.
- The key additional assumption is

$$E(v_{it}|\boldsymbol{x}_i) = 0. \tag{11}$$

. . . This assumption will insure that pooled OLS will be consistent (and unbiased).

. . . But the fact that there is correlation in the errors across observations for the same individual implies that pooled OLS will not be efficient.

The Random Effects Estimator

- The Random Effects (RE) estimator is a GLS estimator, which transforms the data to eliminate the correlation over time in an individual's error terms.
- This is similar to the time series model with AR1 errors, where we used a transformation to eliminate the serial correlation in the Prais-Winsten estimator.
- The **random effects** transformation takes the form

$$\begin{aligned} y_{it} - \theta\bar{y}_i = {} & \beta_0(1-\theta) + \beta_1(x_{it1} - \theta\bar{x}_{i1}) + \cdots + \beta_k(x_{it} - \theta\bar{x}_{i1}) \\ & + v_{it} - \theta\bar{v}_i \end{aligned} \tag{12}$$

or

$$\check{y}_{it} = \beta_0\check{x}_0 + \beta_1\check{x}_{it1} + \cdots + \beta_1\check{x}_{itk} + \check{v}_{it} \tag{13}$$

where $\check{x}_0 = (1-\theta)$, $\check{y}_{it} = y_{it} - \theta\bar{y}_i$ (etc.) and

$$\theta = 1 - \left[\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_a^2}\right]^{\frac{1}{2}} \in [0, 1] \tag{14}$$

The Random Effects Estimator (cont'd)

- The RE transformation is sometimes referred to a **quasi-demeaning**, since it involves subtract a portion of the cross-time means.
- In fact, RE is essentially a compromise between Pooled OLS and the FE estimator, since RE reduces to
  - Pooled OLS if $\theta = 0$ (which occurs if $\sigma_a^2 = 0$)
  - FE if $\theta = 1$ (with $\theta \to 1$ as $\sigma_a^2 \to \infty$).
- One advantage of the RE estimator is that one can include regressors that are constant over time for everyone in the sample.
- However, this comes at the cost of having to assume that the unobserved heterogeneity is uncorrelated with the observed regressors.

  ... This is often an unrealistic assumption.

Example #1: Crime Rates in NY City w/time dummies

	Pooled OLS	FE	FD	RE
lprbarr	-0.512	-0.358	-0.327	-0.369
	(0.137)**	(0.059)**	(0.056)**	(0.065)**
lprbconv	-0.389	-0.283	-0.238	-0.288
	(0.085)**	(0.050)**	(0.040)**	(0.050)**
lprbpris	0.110	-0.181	-0.165	-0.170
	(0.083)	(0.045)**	(0.046)**	(0.045)**
lavgsen	-0.121	-0.004	-0.022	-0.012
	(0.105)	(0.033)	(0.026)	(0.032)
lpolpc	0.283	0.419	0.397	0.400
	(0.139)*	(0.086)**	(0.104)**	(0.094)**
ldensity	0.250	0.372	0.134	0.322
	(0.067)**	(0.396)	(0.417)	(0.057)**
lpctymle	0.053	0.359	0.932	0.115
	(0.132)	(0.537)	(0.630)	(0.148)
_cons	-2.139	-0.756	0.020	-1.456
	(0.986)*	(1.248)	(0.014)	(0.638)*
R^2	[illegible]	[illegible]	[illegible]	
N	630	630	540	630

Example #2: Airline Concentration

	Pooled OLS	FE	FD	RE
concen	0.360	0.169	0.176	0.209
	(0.059)**	(0.049)**	(0.043)**	(0.042)**
ldist	-0.902			-0.852
	(0.272)**			(0.272)**
ldistsq	0.103			0.097
	(0.020)**			(0.020)**
y98	0.021	0.023	-0.039	0.022
	(0.004)**	(0.004)**	(0.006)**	(0.004)**
y99	0.038	0.036	-0.048	0.037
	(0.005)**	(0.005)**	(0.005)**	(0.005)**
y00	0.100	0.098		0.098
	(0.006)**	(0.006)**		(0.006)**
_cons	6.209	4.953	0.061	6.222
	(0.912)**	(0.030)**	(0.003)**	(0.914)**

Random Effects or Fixed Effects?

- As noted earlier, the advantage of the RE model is, if the underlying assumptions are true, then both will be consistent and RE will be more efficient than FE.
- Hausman (1978) proposed a test based on this notion, comparing the difference between the two estimators.
- The RE model is clearly rejected in both of our examples, in both cases with a p-value < 0.01.