

# Lecture 9 - Panel Data

Tuesday, October 12, 2021 2:29 PM



Lecture 9 -  
Panel Data

## Panel Data

---

Mauricio Romero

1

## Panel Data

Introduction

Fixed effects

2

## Panel Data

Introduction

Fixed effects

3

## Repeated observations

- Repeated observations
  - Panel data
  - Time-series cross section data
  - Clustered data, etc
- Dynamics effects (dynamic treatment regimes)
- Identification strategies
  - Fixed effects
  - Difference-in-differences

4

## Group Data

- Panel: observe the same units (individuals, firms, countries, schools, etc.) over several time periods
- Time-series cross section: observe different units across time (e.g., different survey rounds of ENOE)
- Clustered data: Natural grouping in the data (e.g., test score data of students across schools)

## Panel Data

Introduction

Fixed effects

6

## Panel Data

Introduction

Fixed effects

7

## Notation

- Sample of  $i = 1, \dots, N$  units from a population
- Time periods  $t = 1, \dots, T$
- For each  $i$  observe  $(Y_{it}, T_{it})$
- The GDP is:

$$y_{it} = \beta_0 + \beta_1 T_{it} + \alpha_j + u_{it}$$

- Do **not** observe  $\alpha_j$
- Assume  $u_{it}$  is white noise (i.i.d. and mean zero)

8

## OLS

- OLS estimator of Y on T yields

$$\begin{aligned}
 \mathbb{E}\beta_1 &= \frac{\text{cov}(T_{it}, Y_{it})}{V(T_{it})} \\
 &= \frac{\text{cov}(T_{it}, \beta_0 + \beta_1 T_{it} + \alpha_i + u_{it})}{V(T_{it})} \\
 &= \frac{\text{cov}(T_{it}, \beta_0)}{V(T_{it})} + \frac{\text{cov}(T_{it}, \beta_1 T_{it})}{V(T_{it})} + \frac{\text{cov}(T_{it}, \alpha_i)}{V(T_{it})} + \frac{\text{cov}(T_{it}, u_{it})}{V(T_{it})} \\
 &= \underbrace{\frac{\text{cov}(T_{it}, \beta_0)}{V(T_{it})}}_0 + \underbrace{\beta_1 \frac{\text{cov}(T_{it}, T_{it})}{V(T_{it})}}_1 + \underbrace{\frac{\text{cov}(T_{it}, \alpha_i)}{V(T_{it})}}_0 + \underbrace{\frac{\text{cov}(T_{it}, u_{it})}{V(T_{it})}}_0 \\
 &= \beta_1 + \frac{\text{cov}(T_{it}, \alpha_i)}{V(T_{it})}
 \end{aligned}$$

- Omitted variable bias

9

## Simulations!

```

#TRUE MODEL
beta0=1 #intercept
beta1=2 #slope
Nobs=1000 #how many observations?
TimePeriods=5 #how many time periods?
unobserved=runif(Nobs,-1,1) #Create individual unobserved factor
Treatment=sample(c(0,1), size=Nobs*TimePeriods, replace=T) #T-{it}
Data=data.frame(IDObs=rep(1:Nobs, each=TimePeriods),
                Period=rep(1:TimePeriods, Nobs),
                alpha_i=rep(unobserved, each=TimePeriods),
                Treatment=Treatment)
#Let's say high unobservable means you always get treated
#Note this negates the random assignment
Data$Treatment[Data$alpha_i>0.5]=1
#Use runif to generate data
Data$Outcome=beta0+beta1*Data$Treatment+Data$alpha_i+rnorm(Nobs*TimePeriods)
summary(lm(Outcome~Treatment, data=Data)) #seems to be biased
    
```

10

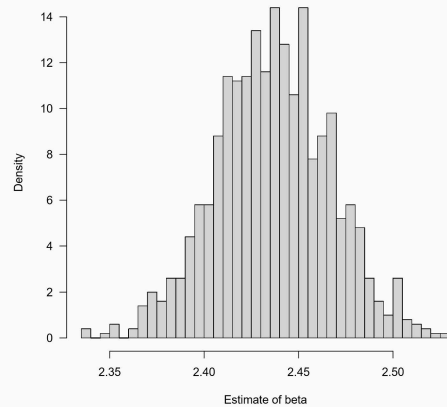
## Simulations!

```

EstimateBeta=NULL
for(r in 1:1000){
  #Use GDP to generate data
  Data$Outcome=beta0+beta1*Data$Treatment+Data$alpha_i+rnorm(Nobs*TimePeriods)
  EstimateBeta=c(EstimateBeta, lm(Outcome~Treatment, data=Data)$coef[2])
}
hist(EstimateBeta, freq=F, breaks=30,
     main="", las=1, xlab="Estimate of beta")
abline(v=beta1, col='red', lwd=3, lty=1)
    
```

11

## Our estimate of the coefficient are biased



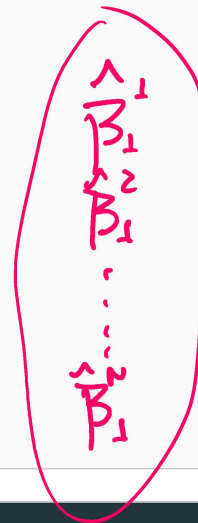
12

## Fixed effects (intuition I)

- Take one  $i$  at a time (like a subset for each  $i$ )

$$\begin{cases} y_{i1} = \beta_0 + \beta_1 T_{i1} + \alpha_i + u_{i1} \\ \vdots \\ y_{iT} = \beta_0 + \beta_1 T_{iT} + \alpha_i + u_{iT} \end{cases}$$

- $\beta_0 + \alpha_i$  is the constant now
- We can estimate  $\beta_1$  for each  $i$  ( $\hat{\beta}_1^i$ )
- Aggregate for all  $i$  to get a more precise estimate of  $\beta_1$
- Maximal precision weighting by  $V(T_{it}|i)$  (why?)



13

## Simulations!

```
#TRUE MODEL
beta0=1 #intercept
beta1=2 #slope
Nobs=1000 #how many observations?
TimePeriods=5 #how many time periods?
unobserved=runif(Nobs,-1,1) #Create individual unobserved factor
Treatment=sample(c(0,1), size=Nobs*TimePeriods, replace=T) #T-{it}
Data=data.frame(IDObs=rep(1:Nobs, each=TimePeriods),
                Period=rep(1:TimePeriods, Nobs),
                alpha_i=rep(unobserved, each=TimePeriods),
                Treatment=Treatment)
#Let's say high unobservable means you always get treated
#Note this negates the random assignment
Data$Treatment[Data$alpha_i>0.5]=1
#Use GDP to generate data
Data$Outcome=beta0+beta1*Data$Treatment+Data$alpha_i+rnorm(Nobs*TimePeriods)
CoeefIndividuals=NULL
VarianceTreatment=NULL
for(i in 1:Nobs){
  CoefIndividuals=c(CoeefIndividuals,
                   lm(Outcome~Treatment, data=Data, subset=IDObs==i)$coef[2])
  VarianceTreatment=c(VarianceTreatment,
                      var(Data$Treatment[which(Data$IDObs==i)]))
}
mean(CoeefIndividuals, na.rm=T)
weighted.mean(CoeefIndividuals, w=VarianceTreatment, na.rm=T)
```

14

## Simulations!

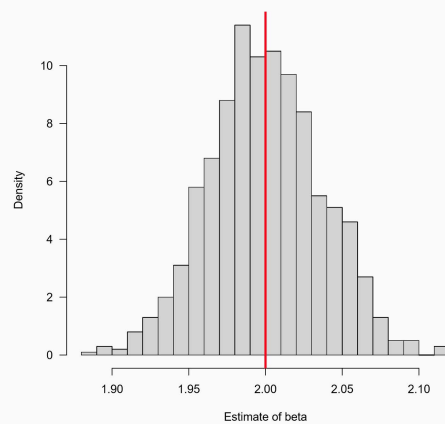
```

EstimateBeta=NULL
for(r in 1:1000){
  #Use GDP to generate data
  Data$Outcome=beta0+beta1*Data$Treatment+Data$alpha_i+rnorm(Nobs*TimePeriods)
  CoefIndividuals=NULL
  VarianceTreatment=NULL
  for(i in 1:Nobs){
    CoefIndividuals=c(CoefIndividuals,
                      lm.fit(y=Data$Outcome[which(Data$IDObs==i)],
                             x=as.matrix(cbind(1,Data$Treatment[which(Data$IDObs==i)]))$coefficients[2]),
                      #lm(Outcome~Treatment, data=Data, subset=IDObs==i)$coef[2])
    VarianceTreatment=c(VarianceTreatment,
                        var(Data$Treatment[which(Data$IDObs==i)]))
  }
  EstimateBeta=c(EstimateBeta,
                 weighted.mean(CoefIndividuals, w=VarianceTreatment, na.rm=T))
}
hist(EstimateBeta, freq=F, breaks=30,
     main="", las=1, xlab="Estimate of beta")
abline(v=beta1, col="red", lwd=3, lty=1)

```

15

## Our estimate of the coefficient are pretty close to the truth



16

## Fixed effects (intuition II)

- Another idea, is to remove the mean for each observation

$$\begin{aligned}
 y_{it} &= \beta_0 + \beta_1 T_{it} + \alpha_i + u_{it} \\
 y_{it} - \bar{Y}_i &= \cancel{\beta_0} + \beta_1 (T_{it} - \bar{T}_i) + \cancel{\alpha_i} + \cancel{u_{it}} - \bar{u}_i \\
 y_{it}^* &= \beta_1 T_{it}^* + u_{it}^*
 \end{aligned}$$

- Estimate via OLS

17

## Simulations!

```
#TRUE MODEL
beta0=1 #intercept
beta1=2 #slope
Nobs=1000 #how many observations?
TimePeriods=5 #how many time periods?
unobserved=runif(Nobs,-1,1) #Create individual unobserved factor
Treatment=sample(c(0,1),size=Nobs*TimePeriods,replace=T) #T-{it}
Data=data.frame(IDObs=rep(1:Nobs,each=TimePeriods),
                Period=rep(1:TimePeriods,Nobs),
                alpha_i=rep(unobserved,each=TimePeriods),
                Treatment=Treatment)
#Let's say high unobservable means you always get treated
#Note this negates the random assignment
Data$Treatment[Data$alpha_i>0.5]=1
#Use GDP to generate data
Data$Outcome=beta0+beta1*Data$Treatment+Data$alpha_i+rnorm(Nobs*TimePeriods)
Data$MeanOutcome <- ave(Data$Outcome, Data$IDObs)
Data$MeanTreatment <- ave(Data$Treatment, Data$IDObs)
Data$OutcomeDemean=Data$Outcome-Data$MeanOutcome
Data$TreatmentDemean=Data$Treatment-Data$MeanTreatment
summary(lm(OutcomeDemean~TreatmentDemean, data=Data))
```

Yic  
Tic

18

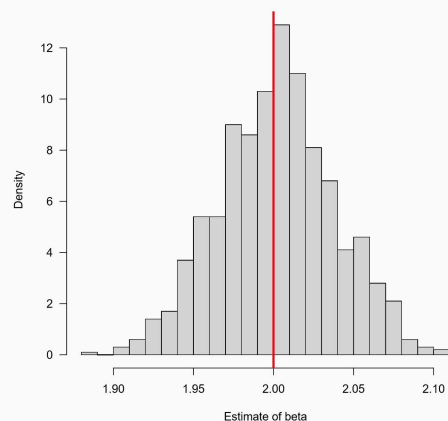
## Simulations!

```
EstimateBeta=NULL
for(r in 1:1000){
  Data$Outcome=beta0+beta1*Data$Treatment+Data$alpha_i+rnorm(Nobs*TimePeriods)
  Data$MeanOutcome <- ave(Data$Outcome, Data$IDObs)
  Data$MeanTreatment <- ave(Data$Treatment, Data$IDObs)
  Data$OutcomeDemean=Data$Outcome-Data$MeanOutcome
  Data$TreatmentDemean=Data$Treatment-Data$MeanTreatment
  EstimateBeta=c(EstimateBeta,
                 lm(OutcomeDemean~TreatmentDemean, data=Data)$coef[2])
}

hist(EstimateBeta, freq=F, breaks=30,
     main="", las=1, xlab="Estimate of beta")
abline(v=beta1, col="red", lwd=3, lty=1)
```

19

## Our estimate of the coefficient are pretty close to the truth



20

## Fixed effects formally

DGP:  $y_{it} = \beta_0 + \beta_1 T_{it} + \alpha_i + u_{it}$

$$y_{it} = \beta_0 + \beta_1 T_{it} + \sum_{j=1}^N 1_{i=j} + u_{it}$$

- $1_{i=j}$  is equal to one if  $j = i$ , zero otherwise — dummy for each group
- By the FWL theorem (decomposition theorem) this is equivalent to:
  - OLS of  $y_{it}$  with respect to  $\sum_{j=1}^N 1_{i=j}$ , and take the residuals ( $\tilde{y}_{it}$ )
  - OLS of  $T_{it}$  with respect to  $\sum_{j=1}^N 1_{i=j}$ , and take the residuals ( $\tilde{T}_{it}$ )
  - OLS of  $\tilde{y}_{it}$  with respect to  $\tilde{T}_{it}$
- OLS with respect to  $\sum_{j=1}^N 1_{i=j}$  equivalent to subtracting group mean
- Same as the transformation we discussed in the previous slide

21

## Fixed effects formally

$$y_{it} = \beta_0 + \beta_1 T_{it} + \sum_{j=1}^N 1_{i=j} + u_{it}$$

- By the FWL theorem:

$$\hat{\beta}_1 = \frac{\sum_i \hat{\beta}_1^i V(T_{it}|i) P(i)}{\sum_i V(T_{it}|i) P(i)}$$

- Same as the estimate from doing OLS one  $i$  at a time

22

## Simulations!

```

beta0=1 #intercept
beta1=2 #slope
Nobs=1000 #how many observations?
TimePeriods=5 #how many time periods?
unobserved=runif(Nobs,-1,1) #Create individual unobserved factor
Treatment=sample(c(0,1), size=Nobs*TimePeriods, replace=T) #T-{it}
Data=data.frame(IDObs=rep(1:Nobs, each=TimePeriods),
                Period=rep(1:TimePeriods, Nobs),
                alpha_i=rep(unobserved, each=TimePeriods),
                Treatment=Treatment)
Data$Treatment[Data$alpha_i>0.5]=1
Data$Outcome=beta0+beta1*Data$Treatment+Data$alpha_i+rnorm(Nobs*TimePeriods)
Data$MeanOutcome <- ave(Data$Outcome, Data$IDObs)
Data$MeanTreatment <- ave(Data$Treatment, Data$IDObs)
Data$OutcomeDemean=Data$Outcome-Data$MeanOutcome
Data$TreatmentDemean=Data$Treatment-Data$MeanTreatment
lm(OutcomeDemean~TreatmentDemean, data=Data)$coef[2]
#felm(Outcome~Treatment|IDObs, data=Data)$coefficients[1]
CoeffIndividuals=NULL
VarianceTreatment=NULL
for(i in 1:Nobs){
  CoeffIndividuals=c(CoeffIndividuals,
                    lm(Outcome~Treatment, data=Data, subset=IDObs==i)$coef[2])
  VarianceTreatment=c(VarianceTreatment,
                      var(Data$Treatment[which(Data$IDObs==i)]))
}
weighted.mean(CoeffIndividuals, w=VarianceTreatment, na.rm=T)
    
```

23



## Fixed effects formally

- Thus, the following are algebraically equivalent:
  - Dummy variable OLS with  $1_{i=j}$
  - Variance weighted average of coefficients from OLS for each  $i$
  - OLS after demeaning (removing  $i$ -specific means)
- This is “one-way fixed effects” regression
- Addresses “time-invariant” confounders

24

## Fixed effects visually (from Nick Huntington-Klein)

<http://nickchk.com/anim/Animation%20of%20Fixed%20Effects.gif>

25

## Fixed effects and causality

- $Y_{1it}$  and  $Y_{0it}$  are period-specific potential outcomes
- $T_{it}$  is the treatment assigned to  $i$  in period  $t$
- We observe

$$Y_{it} = T_{it} Y_{1it} + (1 - T_{it}) Y_{0it}$$

26

## Fixed effects and causality

- Assumption 1:  $T_{it}$  is conditionally mean independent in any given period

$$\mathbb{E}[Y_{0it}|\alpha_i, X_{it}, T_{it}] = \mathbb{E}[Y_{0it}|\alpha_i, X_{it}]$$

- Assumption 2: Linearity

$$\mathbb{E}[Y_{0it}|\alpha_i, X_{it}] = \beta_0 + \alpha_i + X_{it}\gamma$$

- Assumption 3: Constant additive effects:

$$\mathbb{E}[Y_{1it}|\alpha_i, X_{it}] = \mathbb{E}[Y_{0it}|\alpha_i, X_{it}] + \beta_1$$

Under these assumptions, “one-way fixed effects” yields consistent estimator for  $\beta_1$

27

## Problems that fixed effects cannot solve

- Reverse causality
  - Assumption 1 implies potential outcomes uncorrelated with past, current and future treatment!
- Time-varying unobserved heterogeneity

## Regressors that are constant within strata

- If a regressor is constant within an fixed-effect strata, then it is perfectly collinear with that strata dummy
  - A time-invariant regressor in the panel context
- When you fit fixed effects, these strata-invariant regressors must be dropped
- With fixed effects, what matters is whether the demean variables are constant

29

## Clustering standard errors by fixed effect strata

- We cluster to account for dependencies in the treatment
- If treatments are assigned randomly within fixed effect strata (even if treatment probabilities differ across strata), no need to cluster by strata
- If treatment assignment at the strata level (or treatment exhibits positive or negative dependence within a strata), then cluster by strata
- “felm” from the “lfe” package easiest way in R to do fixed effects and cluster

30

## Fixed effects estimators

- Typically we care about  $\beta$ , but unit fixed effects  $\alpha_i$  could be of interest
  - $\hat{\alpha}_i$  from dummy variable regression is unbiased but not consistent for  $\alpha_i$  (based on fixed  $T$  and  $N \rightarrow \infty$ )

## What I will not cover

- Huge literature on panel data
- This is not a review of panel econometrics; for that see Wooldridge and other excellent textbooks
- I won't be covering a lot of it
  - Random effects
  - First difference
  - Arrelano and Bond
  - And much more...
- Goal is to present the modal regression model used in difference-in-differences (next class)