FIN 620
Emp. Methods in Finance

Lecture 4 – Panel Data

Professor Todd Gormley

Announcements

- Option exercise #2 covers today and last class's material

Background readings

- Angrist and Pischke
  - *Sections 5.1, 5.3*
- Wooldridge
  - *Chapter 10 and Sections 13.9.1, 15.8.2, 15.8.3*
- Greene
  - *Chapter 11*

Outline for Today

- Quick review
- Motivate how panel data is helpful
  - Fixed effects model
  - Random effects model
  - First differences
  - Lagged *y* models
- Student presentations of "Causality"

Quick Review *[Part 1]*

- What is the key assumption needed for us to make causal inferences? And what are the ways in which it can be violated?
  - **Answer =** CMI is violated whenever an independent variable, x, is correlated with the error, u. This occurs when there is…
    - Omitted variable bias
    - Measurement error bias
    - Simultaneity bias

Quick Review *[Part 2]*

- When is it possible to determine the sign of an omitted variable bias?

Quick Review *[Part 3]*

- When is measurement error of the <u>dependent</u> variable problematic (for identifying the causal CEF)?

Quick Review *[Part 4]*

- What is the bias on the coefficient of *x*, <u>and</u> on other coefficients when an *indep-endent* variable, *x*, is measured with error?
  - **Answer =** Hard to know!
    - If ME is uncorrelated with observed *x*, no bias
    - If ME is uncorrelated with unobserved *x**, the coefficient on *x* has an attenuation bias, but the sign of the bias on all other coefficients is unclear

Quick Review *[Part 5]*

- When will an estimation suffer from simultaneity bias?

Outline for Panel Data

- Motivate how panel data is helpful
- Fixed effects model
  - Benefits [There are many]
  - Costs [There are some…]
- Random effects model
- First differences
- Lagged *y* models

Motivation *[Part 1]*

- As noted in prior lecture, omitted variables pose a substantial hurdle in our ability to make causal inferences
- What's worse… many of them are inherently unobservable to researchers

Motivation *[Part 2]*

- E.g., consider the firm-level estimation

$$leverage_{i,j,t} = \beta_0 + \beta_1 profit_{i,j,t-1} + u_{i,j,t}$$

where *leverage* is debt/assets for firm *i*, operating in industry *j* in year *t*, and *profit* is the firm's net income/assets

What might be some <u>unobservable</u> omitted variables in this estimation?

Motivation *[Part 3]*

- Oh, there are so, so many…
  - Managerial talent and/or risk aversion
  - Industry supply and/or demand shock
  - Cost of capital
  - Investment opportunities
  - And so on…

Sadly, this is <u>easy</u> to do with other dependent or independent variables…

- Easy to think of ways these might be affect leverage and be correlated with profits

Motivation *[Part 4]*

- Using observations from various geographical regions (e.g., state or country) opens even more possibilities…
  - **Can you think of some unobserved variables that might be related to a firm's location?**
    - **Answer:** any unobserved differences in local economic environment, e.g., institutions, protection of property rights, financial development, investor sentiment, regional demand shocks, etc.

Motivation *[Part 5]*

- Sometimes, we can control for these unobservable variables using proxy variables
  - **But what assumption was required for a proxy variable to provide consistent estimates on the other parameters?**
    - **Answer:** It needs to be a sufficiently good proxy such that the unobserved variable can't be correlated with the other explanatory variables after we control for the proxy variable… **This might be hard to find**

Panel data to the rescue…

- Thankfully, panel data can help us with a particular type of unobserved variable…
  - **What type of unobserved variable does panel data help us with, and why?**
  - **Answer** = It helps us with time-invariant omitted variables; now, let's see why… *[Actually, it helps with any unobserved variable that doesn't vary within groups of observations]*

Outline for Panel Data

- Motivate how panel data is helpful
- Fixed effects model
  - Benefits [There are many]
  - Costs [There are some…]
- Random effects model
- First differences
- Lagged *y* models

Panel data

- **Panel data** = whenever you have multiple observations per unit of observation *i* (e.g., you observe each firm over multiple years)
  - Let's assume *N* units *i*
  - And *T* observations per unit *i [i.e., balanced panel]*
    - **Ex. #1** – You observe 5,000 firms in Compustat over a twenty-year period [i.e., N=5,000, T=20]
    - **Ex. #2** – You observe 1,000 CEOs in Execucomp over a 10-year period [i.e., N=1,000, T=10]

Time-invariant unobserved variable

- Consider the following model…

Unobserved, time-invariant variable, *f*

$$y_{i,t} = \alpha + \beta x_{i,t} + \delta f_i + u_{i,t}$$

where

$$E(u_{i,t}) = 0$$

$$corr(x_{i,t}, f_i) \neq 0$$

$$corr(f_i, u_{i,t}) = 0$$

$$corr(x_{i,t}, u_{i,s}) = 0 \text{ for all } s, t$$

These implies what?
Answer: If don't control for *f*, we will have an omitted variable bias

Note: This is stronger assumption then we usually make; it's called strict exogeneity. In words, this assumption means what?

If we ignore *f*, we get OVB

- If estimate the model…

$$y_{i,t} = \alpha + \beta x_{i,t} + \underbrace{v_{i,t}}_{\delta f_i + u_{i,t}}$$

  - x is correlated with the disturbance v (through its correlation with the unobserved variable, f, which is now part of the disturbance)
  - Easy to show $\hat{\beta}^{OLS} = \beta + \delta \frac{\sigma_{xf}}{\sigma_x^2}$ ← **This is standard OVB… coefficient from regression of omitted var., *f*, on *x* times the true coeff. on *f***

Can solve this by transforming data

- First, notice that if you take the population mean of the dependent variable for each unit of observation, *i*, you get…

$$\bar{y}_i = \alpha + \beta \bar{x}_i + \delta f_i + \bar{u}_i$$

Again, I assumed there are *T* obs. per unit *i*

where

$$\bar{y}_i = \frac{1}{T}\sum_t y_{i,t}, \quad \bar{x}_i = \frac{1}{T}\sum_t x_{i,t}, \quad \bar{u}_i = \frac{1}{T}\sum_t u_{i,t}$$

Transforming data *[Part 2]*

- Now, if we subtract $\bar{y}_i$ from $y_{i,t}$, we have

$$y_{i,t} - \bar{y}_i = \beta\left(x_{i,t} - \bar{x}_i\right) + \left(u_{i,t} - \bar{u}_i\right)$$

  - And look! The unobserved variable, f_i, is gone (as is the constant) because it is time-invariant
  - With our assumption of strict exogeneity earlier, easy to see that $\left(x_{i,t} - \bar{x}_i\right)$ is uncorrelated with the new disturbance, $\left(u_{i,t} - \bar{u}_i\right)$, which means…

?

Fixed Effects (or Within) Estimator

- **Answer:** OLS estimation of transformed model will yield a consistent estimate of β
- The prior transformation is called the **"within transformation**" because it demeans all variables *within* their group
  - In this case, the "group" was each cross-section of observations over time for each firm
  - This is also called the **FE estimator**

Unobserved heterogeneity – *Tangent*

- Unobserved variable, *f*, is very general
  - Doesn't just capture one unobserved variable; captures **all** unobserved variables that don't vary within the group
  - This is why we often just call it **"unobserved heterogeneity"**

FE Estimator – **Practical Advice**

- When you use the fixed effects (FE) estimator in programs like Stata, it does the within transformation for you
- <u>Don't</u> do it on your own because…
  - The degrees of freedom (doF) (which are used to get the standard errors) *sometimes* need to be adjusted down by the number of panels, N
  - What adjustment is necessary depends on whether you cluster, etc.

Least Squares Dummy Variable (LSDV)

- Another way to do the FE estimation is by adding indicator (dummy) variables
  - Notice that the coefficient on f_i, δ, doesn't really have any meaning; so, can just rescale the unobserved f_i to make it equal to 1

$$y_{i,t} = \alpha + \beta x_{i,t} + f_i + u_{i,t}$$

  - Now, to estimate this, we can just treat each f_i as a parameter to be estimated

LSDV continued…

- I.e., create a dummy variable for each group *i*, and add it to the regression
  - This is **least squares dummy variable** model
  - Now, our estimation equation exactly matches the true underlying model

$$y_{i,t} = \alpha + \beta x_{i,t} + f_i + u_{i,t}$$

  - We get consistent estimates and SE that are identical to what we'd get with within estimator

LSDV – **Practical Advice**

- Because the dummy variables will be collinear with the constant, one of them will be dropped in the estimation
  - Therefore, don't try to interpret the intercept; it is just the average y when all the x's are equal to zero for the group corresponding to the dropped dummy variable
  - In **xtreg, fe**, the reported intercept is just average of individual specific intercepts

LSDV *versus* FE *[Part 1]*

- Can show that LSDV and FE are identical, using partial regression results ***[How?]***
  - Remember, to control for some variable z, we can regress y onto both x and z, or we can just partial z out from both y and x before regressing y on x (i.e., regress residuals from regression of y on z onto residual from regression of x on z)
  - The demeaned variables are the residuals from a regression of them onto the group dummies!

LSDV *versus* FE *[Part 2]*

- Reported R^2 will be larger with LSDV
  - All the dummy variables will explain a lot of the variation in *y*, driving up R^2
  - Within R^2 reported for FE estimator just reports what proportion of the *within* variation in *y* that is explained by the *within* variation in *x*
  - **The within R^2 is usually of more interest to us**

R-squared with FE – **Practical Advice**

- The within R^2 is usually of more interest since it describes explanatory power of x's [after partialling out the FE]
  - The get within R^2, use **xtreg, fe**
- Reporting overall adjusted-R^2 is also useful
  - To get overall R^2, use **areg** command instead of **xtreg, fe**. The "overall R^2" reported by **xtreg** does <u>not</u> include variation explained by FE, but the R^2 reported by **areg** does

Outline for Panel Data

- Motivate how panel data is helpful
- Fixed effects model
  - Benefits [There are many]
  - Costs [There are some…]
- Random effects model
- First differences
- Lagged *y* models

FE Estimator – **Benefits** *[Part 1]*

- There are many benefits of FE estimator
  - Allows for arbitrary correlation between each fixed effect, f_i, and each x within group i
    - I.e., its very general and not imposing much structure on what the underlying data must look like
  - Very intuitive interpretation; coefficient is identified using only changes within cross-sections

FE Estimator – **Benefits** *[Part 2]*

- It is also very flexible and can help us control for many types of unobserved heterogeneities
  - Can add year FE if worried about unobserved heterogeneity across time *[e.g., macroeconomic shocks]*
  - Can add CEO FE if worried about unobserved heterogeneity across CEOs *[e.g., talent, risk aversion]*
  - Add industry-by-year FE if worried about unobserved heterogeneity across industries over time *[e.g., investment opportunities, demand shocks]*

FE Estimator – *Tangent [Part 1]*

- FE estimator is very general
  - It applies to any scenario where observations can be grouped together
    - Ex. #1 – Firms can be grouped by industry
    - Ex. #2 – CEOs observations (which may span multiple firms) can be grouped by CEO-firm combinations
  - Textbook example of grouping units i across time is just example (though, the most common)

FE Estimator – *Tangent [Part 2]*

- Once you can construct groups, you can remove any unobserved group-level heterogeneity by adding group FE
  - Consistency just requires there be many groups

Outline for Panel Data

- Motivate how panel data is helpful
- Fixed effects model
  - Benefits [There are many]
  - Costs [There are some…]
- Random effects model
- First differences
- Lagged *y* models

FE Estimator – **Costs**

- However, FE estimator also has its costs
  - Can't identify variables that don't vary within group
  - Subject to potentially large measurement error bias
  - Can be hard to estimate in some cases
  - Miscellaneous issues

<u>**FE Cost #1**</u> – Can’t estimate some var.

- If no within-group variation in the independent variable, x, of interest, can’t disentangle it from group FE
  - It is collinear with group FE; and will be dropped by computer or swept out in the within transformation

FE Cost #1 – *Example*

- Consider following CEO-level estimation

$$\ln(totalpay)_{ijt} = \alpha + \beta_1 \ln(firmsize)_{ijt} + \beta_1 volatility_{ijt}$$
$$+ \beta_3 female_i + \delta_t + f_i + \lambda_j + u_{ijt}$$

  - Ln(*totalpay*) is for CEO *i*, firm *j*, year *t*
  - Estimation includes year, CEO, and firm FE

- **What coefficient can't be estimated?**
  - **Answer:** β_3! Being female probably doesn't vary within the group of each CEO's observations; i.e., it is collinear with the CEO fixed effect

FE Cost #1 – **Practical Advice**

- Be careful of this!
  - Programs like **xtreg** are good about dropping the female variable and not reporting an estimate…
  - But, if you create dummy variables yourself and input them yourself, the estimation might drop one of them rather than the female indicator
    - **I.e., you'll get an estimate for β_3, but it has no meaning! It's just a random intercept value that depends entirely on the random FE dropped by Stata**

FE Cost #1 – Any Solution?

- Instrumental variables can provide a possible solution for this problem
  - See Hausman and Taylor (Econometrica 1981)
  - **We will discuss this next week**

FE Cost #2 – Measurement error *[P1]*

- Measurement error of independent variable (and resulting biases) can be amplified
  - Think of there being two types of variation
    - Good (meaningful) variation
    - Noise variation because we don't perfectly measure the underlying variable of interest
  - Adding FE can sweep out a lot of the good variation; fraction of remaining variation coming from noise goes up ***[What will this do?]***

FE Cost #2 – Measurement error *[P2]*

- **Answer:** Attenuation bias on mismeasured variable will go up!
  - ***Practical advice:*** Be careful in interpreting 'zero' coefficients on potentially mismeasured regressors; might just be attenuation bias!
  - And remember, sign of bias on other coefficients will be generally difficult to know

FE Cost #2 – Measurement error *[P3]*

- Problem can also apply even when all variables are *perfectly* measured ***[How?]***
  - **Answer:** Adding FE might throw out ***relevant*** variation; e.g., y in firm FE model might respond to <u>sustained</u> changes in x, rather than <u>transitory</u> changes *[see McKinnish 2008 for more details]*
  - With FE you'd only have the transitory variation leftover; might find x uncorrelated with y in FE estimation even though sustained changes in x is most important determinant of y

FE Cost #2 – *Example*

- Difficult to identify causal effect of credit shocks on firm output because credit shocks coincide with demand shocks [i.e., OVB]
  - Paravisini, Rappoport, Schnabl, Wolfenzon (2014) used product-level export data & shock to some Peru banks to address this
    - Basically regressed product output on total firm credit, and added firm, bank, and product×destination FE (i.e., dummy for selling a product to a particular country!)
    - Found small effect… **[Concern?]**

FE Cost #2 – *Example* continued

- **Concern** = Credit extended to firms may be measured with error!
  - E.g., some loan originations and payoffs may not be recorded in timely fashion
  - Need to be careful interpreting a coefficient from a model with so many FE as "small"
    - **Note:** This paper is actually very good (and does IV as well), and the authors are very careful to not interpret their findings as evidence that financial constraints only have a "small" effect

FE Cost #2 – Any solution?

- Admittedly, measurement error, in general, is difficult to address
- For examples on how to deal with measurement error, see following papers
  - Griliches and Hausman (*JoE* 1986)
  - Biorn (*Econometric Reviews* 2000)
  - Erickson and Whited *(JPE* 2000, *RFS* 2012*)*
  - Almeida, Campello, and Galvao (*RFS* 2010)

FE Cost #3 – Computation issues *[P1]*

- Estimating a model with multiple types of FE can be computationally difficult
  - When more than one type of FE, you **cannot** remove both using within-transformation
    - Generally, you can only sweep one away with within-transformation; other FE dealt with by adding dummy variable to model
    - E.g., firm and year fixed effects **[See next slide]**

FE Cost #3 – Computation issues *[P2]*

- Consider below model:

Year FE

Firm FE

$$y_{i,t} = \alpha + \beta x_{i,t} + \delta_t + f_i + u_{i,t}$$

  - To estimate this in Stata, we'd use a command something like the following…

xtset *firm*

xi: xtreg *y* *x* i.*year*, fe

Tells Stata that panel dimension is given by *firm* variable

Tells Stata to create and add dummy variables for year variable

Tells Stata to remove FE for panels (i.e., firms) by doing within-transformation

FE Cost #3 – Computation issues *[P3]*

- Dummies not swept away in within-transformation are estimated
  - With year FE, this isn't problem because there aren't that many years of data
  - If had to estimate 1,000s of firm FE, however, it might be a problem
    - In fact, this is why we sweep away the firm FE rather than the year FE; there are more firms!

FE Cost #3 – *Example*

- But computational issues is becoming increasingly more problematic
  - Researchers using larger datasets with many more complicated FE structures
  - E.g., if you try adding both firm and industry×year FE, you'll have a problem
    - Estimating 4-digit SIC×year and firm FE in Compustat requires ≈ 40 GB memory
    - No one had this; hence, no one did it…

FE Cost #3 – Any Solution?

- Yes, there are some potential solutions
  - Gormley and Matsa (2014) discusses some of these solutions in Section 4
  - We will come back to this in "Common Limitations and Errors" lecture

FE – Some Remaining Issues

- Two more issues worth noting about FE
  - Predicted values of unobserved FE
  - Non-linear estimations with FE and the incidental parameter problem

Predicted values of FE *[Part 1]*

- Sometimes, predicted value of unobserved FE is of interest
- Can get predicted value using

$$\hat{f}_i = \bar{y}_i - \hat{\beta}\bar{x}_i, \text{ for all } i = 1,\ldots,N$$

  - E.g., Bertrand and Schoar (QJE 2003) did this to back out CEO fixed effects
    - They show that the CEO FE are jointly statistically significant from zero, suggesting CEOs have 'styles' that affect their firms

Predicted values of FE *[Part 2]*

- But be careful with using these predicted values of the FE
  - They are unbiased, **<u>but</u>** inconsistent
    - As sample size increases (and we get more groups), we have more parameters to estimate… never get the necessary asymptotics
    - We call this the **Incidental Parameters Problem**

Predicted values of FE *[Part 3]*

- Moreover, doing an F-test to show they are statistically different from zero is only valid under rather strong assumptions
  - Need to assume errors, *u*, are distributed normally, homoskedastic, and serially uncorrelated
  - See Wooldridge (2010, Section 10.5.3) and Fee, Hadlock, and Pierce (2011) for more details

Nonlinear models with FE *[Part 1]*

- Because we don't get consistent estimates of the FE, we can't estimate nonlinear panel data models with FE
  - In practice, Logit, Tobit, Probit should **not** be estimated with many fixed effects
  - They only give consistent estimates under rather strong and unrealistic assumptions

Nonlinear models with FE *[Part 2]*

- E.g., Probit with FE requires…
  - Unobserved f_i is to be distributed normally
  - f_i and $x_{i,t}$ to be independent
- And Logit with FE requires…
  - No serial correlation of y after conditioning on the observable x and unobserved f
- For more details, see…
  - Wooldridge (2010), Sections 13.9.1, 15.8.2-3
  - Greene (2004) – uses simulation to show how bad

Why should we believe this to be true?

Almost surely not true in CF

Probably unlikely in many CF settings

Outline for Panel Data

- Motivate how panel data is helpful
- Fixed effects model
  - Benefits [There are many]
  - Costs [There are some…]
- Random effects model
- First differences
- Lagged *y* models

Random effects (RE) model *[Part 1]*

- Very similar model as FE…

$$y_{i,t} = \alpha + \beta x_{i,t} + f_i + u_{i,t}$$

- But one big difference…
  - It assumes that unobserved heterogeneity, f_i, and observed x's are <u>uncorrelated</u>
    - What does this imply about consistency of OLS?
    - Is this a realistic assumption in corporate finance?

Random effects (RE) model *[Part 2]*

- **Answer #1** – That assumption means that OLS would give you consistent estimate of β!
- Then why bother?
  - **Answer**… potential efficiency gain *relative* to FE
    - FE is no longer most efficient estimator. If our assumption is correct, we can get more efficient estimate by not eliminating the FE and doing generalized least squares *[**Note:** can't just do OLS; it will be consistent as well but SE will be wrong since they ignore serial correlation]*

Random effects (RE) model *[Part 3]*

- **Answer #2** – The assumption that f and x are uncorrelated is likely <u>unrealistic</u> in CF
  - The violation of this assumption is whole motivation behind why we do FE estimation!
    - Recall that correlation between unobserved variables, like managerial talent, demand shocks, etc., and x will cause omitted variable bias

Random effects – **My Take**

- In practice, RE model is not very useful
  - As Angrist-Pischke (page 223) write,
    - Relative to fixed effects estimation, random effects requires stronger assumptions to hold
    - Even if right, asymptotic efficiency gain likely modest
    - And finite sample properties can be worse
  - **Bottom line, don't bother with it**

Outline for Panel Data

- Motivate how panel data is helpful
- Fixed effects model
  - Benefits [There are many]
  - Costs [There are some…]
- Random effects model
- First differences
- Lagged *y* models

First differencing (FD) *[Part 1]*

- First differencing is another way to remove unobserved heterogeneities
  - Rather than subtracting off the group mean of the variable from each variable, you instead subtract the lagged observation
  - Easy to see why this also works…

First differencing (FD) *[Part 2]*

- Notice that, $y_{i,t} = \alpha + \beta x_{i,t} + f_i + u_{i,t}$

  $y_{i,t-1} = \alpha + \beta x_{i,t-1} + f_i + u_{i,t-1}$

- From this, we can see that

  $$y_{i,t} - y_{i,t-1} = \beta\left(x_{i,t} - x_{i,t-1}\right) + \left(u_{i,t} - u_{i,t-1}\right)$$

  - When will OLS estimate of this provide a consistent estimate of β?
    - **Answer:** With same strict exogeneity assumption of FE (i.e., $x_{i,t}$ and $u_{i,s}$ are uncorrelated for all t and s)

Note: we'll lose on observation per cross-section because there won't be a lag

First differences (without time)

- First differences can also be done even when observations within groups aren't ordered by time
  - Just order the data within groups in whatever way you want, and take 'differences'
  - Works, but admittedly, not usually done

FD *versus* FE *[Part 1]*

- When just two observations per group, they are identical to each other
- In other cases, both are consistent; difference is generally about efficiency
  - FE is more efficient if disturbances, $u_{i,t}$, are serially uncorrelated
  - FD is more efficient if disturbance, $u_{i,t}$, follow a random walk

Which is true?
Unclear. Truth is that it is probably something in between

FD *versus* FE *[Part 2]*

- If strict exogeneity is violated *(i.e., $x_{i,t}$ is correlated with $u_{i,s}$ for $s \neq t$)*, FE might be better
  - As long as we believe $x_{i,t}$ and $u_{i,t}$ are uncorrelated, the FE's inconsistency shrinks to 0 at rate 1/T; but FD gets no better with larger T
  - **Remember:** T is the # of observations per group
- But, if y and x are spuriously correlated, and N is small, T large, FE can be quite bad

FD *versus* FE *[Part 3]*

- Bottom line: not a bad idea to try both…
  - If different, you should try to understand why
  - With an omitted variable or measurement error, you'll get diff. answers with FD and FE
    - In fact, Griliches and Hausman (1986) shows that because measurement error causes predictably different biases in FD and FE, you can (under certain circumstances) use the biased estimates to back out the true parameter

Outline for Panel Data

- Motivate how panel data is helpful
- Fixed effects model
  - Benefits [There are many]
  - Costs [There are some…]
- Random effects model
- First differences
- Lagged *y* models

Lagged dependent variables with FE

- We cannot easily estimate models with both a lagged dep. var. **<u>and</u>** unobserved FE

$$y_{i,t} = \alpha + \rho y_{i,t-1} + \beta x_{i,t} + f_i + u_{i,t}, \quad |\rho| < 1$$

  - Same as before, but now true model contains lagged *y* as independent variable
    - Can't estimate with OLS <u>even</u> if *x* & *f* are uncorrelated
    - Can't estimate with FE

Lagged *y* & FE – Problem with OLS

- To see the problem with OLS, suppose you estimate the following:

$$y_{i,t} = \alpha + \rho y_{i,t-1} + \beta x_{i,t} + \underbrace{v_{i,t}}_{f_i + u_{i,t}}$$

  - But, $y_{i,t-1} = \alpha + \rho y_{i,t-2} + \beta x_{i,t-1} + f_i + u_{i,t-1}$
  - Thus, $y_{i,t-1}$ and composite error, $v_{i,t}$ are positively correlated because they both contain f_i
  - I.e., you get omitted variable bias

Lagged *y* & FE – Problem with FE

- Will skip the math, but it is always biased
  - Basic idea is that if you do a within transformation, the lagged mean of *y*, which will be on RHS of the model now, will always be negatively correlated with demeaned error, *u*
    - **Note #1 –** This is true even if there was no unobserved heterogeneity, *f*; FE with lagged values is always bad idea
    - **Note #2:** Same problem applies to FD
  - Problem, however, goes away as T goes to infinity

How do we estimate this? IV?

- Basically, you're going to need instrument; we will come back to this next week….

Lagged *y versus* FE – Bracketing

- Suppose you don't know which is correct
  - Lagged value model: $y_{i,t} = \alpha + \gamma y_{i,t-1} + \beta x_{i,t} + u_{i,t}$
  - Or FE model: $y_{i,t} = \alpha + \beta x_{i,t} + f_i + u_{i,t}$
- Can show that estimate of β>0 will…
  - Be too high if lagged model is correct, but you incorrectly use FE model
  - Be too low if FE model is correct, but you incorrectly used lagged model

Bracketing continued…

- Use this to 'bracket' where true β is…
  - But sometimes, you won't observe bracketing
  - Likely means your model is incorrect in other ways, or there is some severe finite sample bias

Summary of Today *[Part 1]*

- Panel data allows us to control for certain types of unobserved variables
  - FE estimator can control for these potential unobserved variables in very flexible way
  - Greatly reduces the scope for potential omitted variable biases we need to worry about
  - Random effects model is useless in most empirical corporate finance settings

Summary of Today *[Part 2]*

- FE estimator, however, has weaknesses
  - Can't estimate variables that don't vary within groups *[or at least, not without an instrument]*
  - Could amplify any measurement error
    - For this reason, be cautious interpreting zero or small coefficients on possibly mismeasured variables
  - Can't be used in models with lagged values of the dependent variable *[or at least, not without an IV]*

Summary of Today *[Part 3]*

- FE are generally not a good idea when estimating nonlinear models [e.g., Probit, Tobit, Logit]; estimates are **inconsistent**
- First differences can also remove unobserved heterogeneity
  - Largely just differs from FE in terms of relative efficiency; which depends on error structure

In First Half of Next Class

- Instrumental variables
  - What are the necessary assumptions? [E.g., what is the exclusion restriction?]
  - Is there are way we can test whether our instruments are okay?
- Related readings… see syllabus

Assign papers for next week…

- Khwaja and Mian (AER 2008)
  - Bank liquidity shocks
- Paravisini, et al. (ReStud 2014)
  - Impact of credit supply on trade
- Becker, Ivkovic, and Weisbenner (JF 2011)
  - Local dividend clienteles

Break Time

- Let’s take our 10-minute break
- We’ll do presentations when we get back