Causal Inference

Mixtape Session

MIXTAPE SESSIONS

Roadmap

Twoway fixed effects

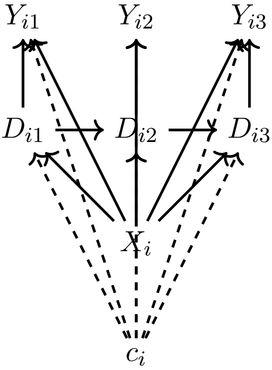
- When working with panel data, the so-called “twoway fixed effects” (TWFE) estimator is the workhorse estimator
- It’s easy to run, a version of OLS, and many people are just interested in mean effects anyway
- It’s the most common model for estimating treatment effects in a difference-in-differences, and so for all these reasons, we need to spend some time understanding what it is

Panel Data

- Panel data: we observe the same units (individuals, firms, countries, schools, etc.) over several time periods
- Often our outcome variable depends on unobserved factors which are also correlated with our explanatory variable of interest
- If these omitted variables are constant over time, we can use panel data estimators to consistently estimate the effect of our explanatory variable

What I will cover

- I will cover pooled OLS and twoway fixed effects
- But I won't be covering random effects, Arrelano and Bond and any number of important panel estimators because the purpose here is to present the modal regression model used in difference-in-differences



Sorry - drawing the DAG for a simple panel model is somewhat messy!

When to use this

- Traditionally, this was used for estimating constant treatment effects with unobserved time-invariant heterogeneity – recall the c_i was constant across all time periods
- It's a linear model, so you'll be estimating conditional mean treatment effects – if you want the median, you can't use this
- Once you enter into a world with dynamic treatment effects and differential timing, this loses all value

Problems that fixed effects cannot solve

- Reverse causality: Becker predicted police reduce crime, but when you regress crime onto police, it's usually positive
  - $\rightarrow$ $\widehat{\beta}_{FE}$ inconsistent unless strict exogeneity conditional on c_i holds
    - $E[\varepsilon_{it}|x_{i1}, x_{i2}, \dots, x_{iT}, c_i] = 0; t = 1, 2, \dots, T$
    - implies ε_{it} uncorrelated with past, current and future regressors
- Time-varying unobserved heterogeneity
  - $\rightarrow$ It's the time-varying unobservables you have to worry about in fixed effects
  - $\rightarrow$ Can include time-varying controls, but as always, don't condition on a collider

Formal panel notation

- Let y and $x \equiv (x_1, x_2, \ldots, x_k)$ be observable random variables and c be an unobservable random variable
- We are interested in the partial effects of variable x_j in the population regression function

$$E[y|x_1, x_2, \ldots, x_k, c]$$

Formal panel notation cont.

- We observe a sample of $i = 1, 2, \ldots, N$ cross-sectional units for $t = 1, 2, \ldots, T$ time periods (a balanced panel)
  - $\rightarrow$ For each unit i, we denote the observable variables for all time periods as $\{(y_{it}, x_{it}) : t = 1, 2, \ldots, T\}$
  - $\rightarrow$ $x_{it} \equiv (x_{it1}, x_{it2}, \ldots, x_{itk})$ is a $1 \times K$ vector
- Typically assume that cross-sectional units are i.i.d. draws from the population: $\{y_i, x_i, c_i\}_{i=1}^{N} \sim i.i.d.$ (cross-sectional independence)
  - $\rightarrow$ $y_i \equiv (y_{i1}, y_{i2}, \ldots, y_{iT})'$ and $x_i \equiv (x_{i1}, x_{i2}, \ldots, x_{iT})$
  - $\rightarrow$ Consider asymptotic properties with T fixed and $N \rightarrow \infty$

Formal panel notation

Single unit:

$$y_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{it} \\ \vdots \\ y_{iT} \end{pmatrix}_{T\times 1} \qquad X_i = \begin{pmatrix} X_{i,1,1} & X_{i,1,2} & X_{i,1,j} & \dots & X_{i,1,K} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{i,t,1} & X_{i,t,2} & X_{i,t,j} & \dots & X_{i,t,K} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{i,T,1} & X_{i,T,2} & X_{i,T,j} & \dots & X_{i,T,K} \end{pmatrix}_{T\times K}$$

Panel with all units:

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{pmatrix}_{NT\times 1} \qquad X = \begin{pmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{pmatrix}_{NT\times K}$$

Unobserved heterogeneity

- For a randomly drawn cross-sectional unit i, the model is given by

$$y_{it} = x_{it}\beta + c_i + \varepsilon_{it},\ t = 1, 2, \ldots, T$$

  - → y_{it}: log wages i in year t
  - → $x_{it} : 1 \times K$ vector of variable events for person i in year t, such as education, marriage, etc. plus an intercept
  - → $\beta : K \times 1$ vector of marginal effects of events
  - → c_i: sum of all time-invariant inputs known to people i (but unobserved for the researcher), e.g., ability, beauty, grit, etc., often called unobserved heterogeneity or fixed effect
  - → ε_{it}: time-varying unobserved factors, such as a recession, unknown to the farmer at the time the decision on the events x_{it} are made, sometimes called idiosyncratic error

Pooled OLS

- When we ignore the panel structure and regress y_{it} on x_{it} we get

$$y_{it} = x_{it}\beta + v_{it};\ t = 1, 2, \ldots, T$$

  with composite error $v_{it} \equiv c_i + \varepsilon_{it}$
- What happens when we regress y_{it} on x_{it} if x is correlated with c_i?
- Then x ends up correlated with v, the composite error term.
- Somehow we need to eliminate this bias, but how?

Pooled OLS

- Main assumption to obtain consistent estimates for β is:
  - → $E[v_{it}|x_{i1}, x_{i2}, \ldots, x_{iT}] = E[v_{it}|x_{it}] = 0$ for $t = 1, 2, \ldots, T$
    - x_{it} are strictly exogenous: the composite error v_{it} in each time period is uncorrelated with the past, current and future regressors
    - But: education x_{it} likely depends on grit and ability c_i and so we have omitted variable bias and $\widehat{\beta}$ is not consistent
  - → No correlation between x_{it} and v_{it} implies no correlation between unobserved effect c_i and x_{it} for all t
    - Violations are common: whenever we omit a time-constant variable that is correlated with the regressors (heterogeneity bias)
  - → Additional problem: v_{it} are serially correlated for same i since c_i is present in each t and thus pooled OLS standard errors are invalid

Pooled OLS

- Always ask: is there a time-constant unobserved variable (c_i) that is correlated with the regressors?
- If yes, then pooled OLS is problematic
- This is how we motivate a fixed effects model: because we believe unobserved heterogeneity is the main driving force making the treatment variable endogenous

Fixed effect regression

- Our unobserved effects model is:

$$y_{it} = x_{it}\beta + c_i + \varepsilon_{it}; t = 1, 2, \ldots, T$$

- If we have data on multiple time periods, we can think of c_i as **fixed effects** to be estimated
- OLS estimation with fixed effects yields

$$(\widehat{\beta}, \widehat{c}_1, \ldots, \widehat{c}_N) = \underset{b, m_1, \ldots, m_N}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it} - x_{it} b - m_i)^2$$

this amounts to including N individual dummies in regression of y_{it} on x_{it}

Derivation: fixed effects regression

$$(\widehat{\beta}, \widehat{c}_1, \ldots, \widehat{c}_N) = \underset{b, m_1, \ldots, m_N}{argmin} \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it} - x_{it} b - m_i)^2$$

The first-order conditions (FOC) for this minimization problem are:

$$\sum_{i=1}^{N} \sum_{t=1}^{T} x'_{it} (y_{it} - x_{it} \widehat{\beta} - \widehat{c}_i) = 0$$

and

$$\sum_{t=1}^{T} (y_{it} - x_{it} \widehat{\beta} - \widehat{c}_i) = 0$$

for $i = 1, \ldots, N$.

Derivation: fixed effects regression

Therefore, for $i = 1, \ldots, N$,

$$\widehat{c}_i = \frac{1}{T} \sum_{t=1}^{T} (y_{it} - x_{it}\widehat{\beta}) = \bar{y}_i - \bar{x}_i \widehat{\beta},$$

where

$$\bar{x}_i \equiv \frac{1}{T} \sum_{t=1}^{T} x_{it}; \bar{y}_i \equiv \frac{1}{T} \sum_{t=1}^{T} y_{it}$$

Plug this result into the first FOC to obtain:

$$\widehat{\beta} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)'(x_{it} - \bar{x}_i) \right)^{-1} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)'(y_{it} - \bar{y}) \right)$$

$$\widehat{\beta} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \ddot{x}_{it}' \ddot{x}_{it} \right)^{-1} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \ddot{x}_{it}' \ddot{y}_{it} \right)$$

with time-demeaned variables $\ddot{x}_{it} \equiv x_{it} - \bar{x}, \ddot{y}_{it} \equiv y_{it} - \bar{y}_i$

Fixed effects regression

Running a regression with the time-demeaned variables $\ddot{y}_{it} \equiv y_{it} - \bar{y}_i$ and $\ddot{x}_{it} \equiv x_{it} - \bar{x}$ is numerically equivalent to a regression of y_{it} on x_{it} and unit specific dummy variables.

Even better, the regression with the time demeaned variables is consistent for β even when $Cov[x_{it}, c_i] \neq 0$ because time-demeaning eliminates the unobserved effects

$$\begin{aligned} y_{it} &= x_{it}\beta + c_i + \varepsilon_{it} \\ \bar{y}_i &= \bar{x}_i\beta + c_i + \bar{\varepsilon}_i \end{aligned}$$

$$\begin{aligned} (y_{it} - \bar{y}_i) &= (x_{it} - \bar{x})\beta + (c_i - c_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \\ \ddot{y}_{it} &= \ddot{x}_{it}\beta + \ddot{\varepsilon}_{it} \end{aligned}$$

Fixed effects regression: main results

- Identification assumptions:
  1. $E[\varepsilon_{it}|x_{i1}, x + i2, \ldots, x_{iT}, c_i] = 0; t = 1, 2, \ldots, T$
     - regressors are strictly exogenous conditional on the unobserved effect
     - allows x_{it} to be arbitrarily related to c_i
  2. $rank\left(\sum_{t=1}^{T} E[\ddot{x}'_{it}\ddot{x}_{it}] \right) = K$
     - regressors vary over time for at least some i and not collinear
- Fixed effects estimator
  1. Demean and regress $\ddot{y}_{it}$ on $\ddot{x}_{it}$ (need to correct degrees of freedom)
  2. Regress y_{it} on x_{it} and unit dummies (dummy variable regression)
  3. Regress y_{it} on x_{it} with canned fixed effects routine
     - Stata: `xtreg y x, fe i(PanelID)`

FE main results

- Properties (under assumptions 1-2):
  - $\rightarrow$ $\widehat{\beta}_{FE}$ is consistent: $\underset{N\to\infty}{plim}\,\widehat{\beta}_{FE,N} = \beta$
  - $\rightarrow$ $\widehat{\beta}_{FE}$ is unbiased conditional on **X**

Fixed effects regression: main issues

- Inference:
  - $\rightarrow$ Standard errors have to be "clustered" by panel unit (e.g., farm) to allow correlation in the ε_{it}'s for the same i.
  - $\rightarrow$ Yields valid inference as long as number of clusters is reasonably large
- Typically we care about β, but unit fixed effects c_i could be of interest
  - $\rightarrow$ $\widehat{c}_i$ from dummy variable regression is unbiased but not consistent for c_i (based on fixed T and $N \rightarrow \infty$)

Application: SASP

- From 2008-2009, I fielded a survey of Internet sex workers (685 respondents, 5% response rate)
- I asked two types of questions: static provider-specific information (e.g., age, weight) and dynamic session information over last 5 sessions
- Let's look at the panel aspect of this analysis together

Risk premium equation

$$
\begin{aligned}
Y_{is} &= \beta_i X_i + \delta D_{is} + \gamma_{is} Z_{is} + u_i + \varepsilon_{is} \\
\ddot{Y}_{is} &= \gamma_{is} \ddot{Z}_{is} + \ddot{\eta}_{is}
\end{aligned}
$$

where Y is log price, D is unprotected sex with a client in a session, X are client and session characteristics, Z is unobserved heterogeneity, and u_i is both unobserved and correlated with Z_{is}.

Table: POLS, FE and Demeaned OLS Estimates of the Determinants of Log Hourly Price for a Panel of Sex Workers

Depvar:	**POLS**	**FE**	**Demeaned OLS**
Unprotected sex with client of any kind	0.013	0.051*	0.051*
	(0.028)	(0.028)	(0.026)
Ln(Length)	-0.308***	-0.435***	-0.435***
	(0.028)	(0.024)	(0.019)
Client was a Regular	-0.047*	-0.037**	-0.037**
	(0.028)	(0.019)	(0.017)
Age of Client	-0.001	0.002	0.002
	(0.009)	(0.007)	(0.006)
Age of Client Squared	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Client Attractiveness (Scale of 1 to 10)	0.020***	0.006	0.006
	(0.007)	(0.006)	(0.005)
Second Provider Involved	0.055	0.113*	0.113*
	(0.067)	(0.060)	(0.048)
Asian Client	-0.014	-0.010	-0.010
	(0.049)	(0.034)	(0.030)
Black Client	0.092	0.027	0.027
	(0.073)	(0.042)	(0.037)
Hispanic Client	0.052	-0.062	-0.062
	(0.080)	(0.052)	(0.045)
Other Ethnicity Client	0.156**	0.142***	0.142***
	(0.068)	(0.049)	(0.045)
Met Client in Hotel	0.133***	0.052*	0.052*
	(0.029)	(0.027)	(0.024)
Gave Client a Massage	-0.134***	-0.001	-0.001
	(0.029)	(0.028)	(0.024)
Age of provider	0.003	0.000	0.000
	(0.012)	(.)	(.)

Table: POLS, FE and Demeaned OLS Estimates of the Determinants of Log Hourly Price for a Panel of Sex Workers

Depvar:	**POLS**	**FE**	**Demeaned OLS**
Body Mass Index	-0.022***	0.000	0.000
	(0.002)	(.)	(.)
Hispanic	-0.226***	0.000	0.000
	(0.082)	(.)	(.)
Black	0.028	0.000	0.000
	(0.064)	(.)	(.)
Other	-0.112	0.000	0.000
	(0.077)	(.)	(.)
Asian	0.086	0.000	0.000
	(0.158)	(.)	(.)
Imputed Years of Schooling	0.020**	0.000	0.000
	(0.010)	(.)	(.)
Cohabitating (living with a partner) but unmarried	-0.054	0.000	0.000
	(0.036)	(.)	(.)
Currently married and living with your spouse	0.005	0.000	0.000
	(0.043)	(.)	(.)
Divorced and not remarried	-0.021	0.000	0.000
	(0.038)	(.)	(.)
Married but not currently living with your spouse	-0.056	0.000	0.000
	(0.059)	(.)	(.)
N	1,028	1,028	1,028
Mean of dependent variable	5.57	5.57	0.00

Heteroskedastic robust standard errors in parenthesis clustered at the provider level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Unit specific time trends often eliminate “results”

Table: Demeaned OLS Estimates of the Determinants of Log Hourly Price for a Panel of Sex Workers with provider specific trends

Depvar:	**FE w/provider trends**
Unprotected sex with client of any kind	0.004
	(0.046)
Ln(Length)	-0.450***
	(0.020)
Client was a Regular	-0.071**
	(0.023)
Age of Client	0.008
	(0.005)
Age of Client Squared	-0.000
	(0.000)
Client Attractiveness (Scale of 1 to 10)	0.003
	(0.003)
Second Provider Involved	0.126*
	(0.055)
Asian Client	-0.048***
	(0.007)
Black Client	0.017
	(0.043)
Hispanic Client	-0.015
	(0.022)
Other Ethnicity Client	0.135***
	(0.031)
Met Client in Hotel	0.073***

Concluding remarks

- This is not a review of panel econometrics; for that see Wooldridge and other excellent options
- We reviewed POLS and TWFE because they are commonly used with individual level panel data and difference-in-differences
- Their main value is how they control for unobserved heterogeneity through a simple demeaning
- Now let's discuss difference-in-differences which will at various times use the TWFE model