Lecture 12:
Panel data II

PPHA 34600

Prof. Fiona Burlig

Harris School of Public Policy
University of Chicago

From last time: regression-based DD

The simplest implementation of DD is just:

$$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$$

From last time: regression-based DD

The simplest implementation of DD is just:

$$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$$

We can implement this via the following regression:

$$Y_i = \alpha + \tau Treat \times Post_{it} + \beta Treat_i + \delta Post_t + \varepsilon_{it}$$

From last time: regression-based DD

The simplest implementation of DD is just:

$$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$$

We can implement this via the following regression:

$$Y_i = \alpha + \tau Treat \times Post_{it} + \beta Treat_i + \delta Post_t + \varepsilon_{it}$$

To link these together, see:
$\bar{Y}(treat, post) = \hat{\alpha} + \hat{\tau} + \hat{\beta} + \hat{\delta}$
$\bar{Y}(treat, pre) = \hat{\alpha} + \hat{\beta}$
$\rightarrow \bar{Y}(treat, post) - \bar{Y}(treat, pre) = \hat{\delta} + \hat{\tau}$

From last time: regression-based DD

The simplest implementation of DD is just:

$$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$$

We can implement this via the following regression:

$$Y_i = \alpha + \tau Treat \times Post_{it} + \beta Treat_i + \delta Post_t + \varepsilon_{it}$$

To link these together, see:

$\bar{Y}(treat, post) = \hat{\alpha} + \hat{\tau} + \hat{\beta} + \hat{\delta}$

$\bar{Y}(treat, pre) = \hat{\alpha} + \hat{\beta}$

$\rightarrow \bar{Y}(treat, post) - \bar{Y}(treat, pre) = \hat{\delta} + \hat{\tau}$

and

$\bar{Y}(untreat, post) = \hat{\alpha} + \hat{\delta}$

$\bar{Y}(untreat, pre) = \hat{\alpha}$

$\rightarrow \bar{Y}(untreat, post) - \bar{Y}(untreat, pre) = \hat{\delta}$

From last time: regression-based DD

The simplest implementation of DD is just:

$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$

We can implement this via the following regression:

$$Y_i = \alpha + \tau Treat \times Post_{it} + \beta Treat_i + \delta Post_t + \varepsilon_{it}$$

To link these together, see:
$\bar{Y}(treat, post) = \hat{\alpha} + \hat{\tau} + \hat{\beta} + \hat{\delta}$
$\bar{Y}(treat, pre) = \hat{\alpha} + \hat{\beta}$
$\rightarrow \bar{Y}(treat, post) - \bar{Y}(treat, pre) = \hat{\delta} + \hat{\tau}$
and
$\bar{Y}(untreat, post) = \hat{\alpha} + \hat{\delta}$
$\bar{Y}(untreat, pre) = \hat{\alpha}$
$\rightarrow \bar{Y}(untreat, post) - \bar{Y}(untreat, pre) = \hat{\delta}$

Running this regression yields $\hat{\tau} = \hat{\tau}^{DD}$

Assessing the validity of the identifying assumption

The identifying assumption for the DD:

In words:

- Parallel counterfactual trends

Assessing the validity of the identifying assumption

The identifying assumption for the DD:

In words:

- Parallel counterfactual trends

In mathier words:

- $E[\varepsilon_{it} | Treat_i, Post_t, X_{it}] = 0$

In different words:

- Conditional on covariates, treatment is as good as randomly assigned

In other different words:

- Treated and untreated units would be on similar trajectories if not for treatment

Assessing the validity of the identifying assumption

The identifying assumption for the DD:

In words:

- Parallel counterfactual trends

In mathier words:

- $E[\varepsilon_{it} | Treat_i, Post_t, X_{it}] = 0$

In different words:

- Conditional on covariates, treatment is as good as randomly assigned

In other different words:

- Treated and untreated units would be on similar trajectories if not for treatment

→ Just controlling for $Treat_i$ and $Post_t$ is usually not good enough!

→ We also need a story for why $Treat_i \times Post_t$ is quasi-random!

Fixed effects models

DD is a specific case of the fixed effects model, which lets us:

- Easily incorporate more units and time periods
- Allow for different units to be treated at different time periods
- Allow the effect of treatment to vary over time
- Enable us to more easily asssess the identifying assumptions

Fixed effects models

Consider a basic panel data regression:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \varepsilon_{it}$$

Fixed effects models

Consider a basic panel data regression:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \varepsilon_{it}$$

We can split ε_{it} into three components:

$$\varepsilon_{it} = \alpha_i + \delta_t + \nu_{it}$$

- α_i: Individual-specific time-invariant bit
- δ_t: Common time-period-specific bit
- ν_{it}: Remaining individual-by-time varying unobserved bit

Fixed effects models

Consider a basic panel data regression:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \varepsilon_{it}$$

We can split ε_{it} into three components:

$$\varepsilon_{it} = \alpha_i + \delta_t + \nu_{it}$$

- α_i: Individual-specific time-invariant bit
- δ_t: Common time-period-specific bit
- ν_{it}: Remaining individual-by-time varying unobserved bit

Can we do better than leaving all of these in the error term?

Fixed effects models

Enter the **fixed effects model**:

- We want to control for α_i and δ_t: they might be correlated with D_{it}

Fixed effects models

Enter the **fixed effects model**:

- We want to control for α_i and δ_t: they might be correlated with D_{it}

There are two main ways to do this:

1. Use dummy variables
2. "De-meaning" the data

Fixed effects models

Enter the **fixed effects model**:

- We want to control for α_i and δ_t: they might be correlated with D_{it}

There are two main ways to do this:

1. Use dummy variables
2. "De-meaning" the data

→ These are both "fixed effects" estimators

Fixed effects estimator 1: Use dummy variables

Consider the following regression model with only individual (no time) effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \nu_{it}$$

Fixed effects estimator 1: Use dummy variables

Consider the following regression model with only individual (no time) effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \nu_{it}$$

Enter the dummy variable approach:

- We'd like to "control for" α_i (ie, separate α_i from the error term)
- But what are these α_is?
- Just individual-specific effects (intercepts, if you want)

Fixed effects estimator 1: Use dummy variables

Consider the following regression model with only individual (no time) effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \nu_{it}$$

Enter the dummy variable approach:

- We'd like to "control for" α_i (ie, separate α_i from the error term)
- But what are these α_is?
- Just individual-specific effects (intercepts, if you want)
- So we should be able to just add a control variable for each person
- We do this with **dummy variables**: $I_i = 1$ for unit i, 0 for all $j \neq i$

Fixed effects estimator 1: Use dummy variables

Consider the following regression model with only individual (no time) effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \nu_{it}$$

Enter the dummy variable approach:

- We'd like to "control for" α_i (ie, separate α_i from the error term)
- But what are these α_is?
- Just individual-specific effects (intercepts, if you want)
- So we should be able to just add a control variable for each person
- We do this with **dummy variables**: $I_i = 1$ for unit i, 0 for all $j \neq i$

$$Y_{it} = X_{it}\beta + \tau D_{it} + \sum_{i=1}^{N} \mathbf{1}[unit = i]_i + \nu_{it}$$

Fixed effects estimator 2: De-meaning

Consider the following regression model with only individual (no time) effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \nu_{it}$$

We can get the same effect as adding dummies by "de-meaning" the data

- In general, define $\tilde{V}_{it} = V_{it} - \frac{1}{T}\sum_{t=1}^{T} V_{it} = V_{it} - \bar{V}_i$

Fixed effects estimator 2: De-meaning

Consider the following regression model with only individual (no time) effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \nu_{it}$$

We can get the same effect as adding dummies by "de-meaning" the data

- In general, define $\tilde{V}_{it} = V_{it} - \frac{1}{T}\sum_{t=1}^{T} V_{it} = V_{it} - \bar{V}_i$
- $\tilde{Y}_{it} = Y_{it} - \bar{Y}_i$: Outcome minus averaged outcome
- $\tilde{X}_{it} = X_{it} - \bar{X}_i$: Covariate(s) minus averaged covariate(s)
- $\tilde{D}_{it} = D_{it} - \bar{D}_i$: Treatment minus averaged treatment
- $\tilde{\alpha}_i = \alpha_i - \bar{\alpha}_i = 0$: Individual effect minus averaged individual effect
- $\tilde{\nu}_{it} = \nu_{it} - \bar{\nu}_i$: Error minus averaged error

Fixed effects estimator 2: De-meaning

Consider the following regression model with only individual (no time) effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \nu_{it}$$

We can get the same effect as adding dummies by "de-meaning" the data

- In general, define $\tilde{V}_{it} = V_{it} - \frac{1}{T}\sum_{t=1}^{T} V_{it} = V_{it} - \bar{V}_i$
- $\tilde{Y}_{it} = Y_{it} - \bar{Y}_i$: Outcome minus averaged outcome
- $\tilde{X}_{it} = X_{it} - \bar{X}_i$: Covariate(s) minus averaged covariate(s)
- $\tilde{D}_{it} = D_{it} - \bar{D}_i$: Treatment minus averaged treatment
- $\tilde{\alpha}_i = \alpha_i - \bar{\alpha}_i = 0$: Individual effect minus averaged individual effect
- $\tilde{\nu}_{it} = \nu_{it} - \bar{\nu}_i$: Error minus averaged error

We can then express the "de-meaned" version of our model:

$$\tilde{Y}_{it} = \tilde{X}_{it}\beta + \tau\tilde{D}_{it} + \tilde{\nu}_{it}$$

Fixed effects estimator 2: De-meaning

Consider the following regression model with only individual (no time) effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \nu_{it}$$

We can get the same effect as adding dummies by "de-meaning" the data

- In general, define $\tilde{V}_{it} = V_{it} - \frac{1}{T}\sum_{t=1}^{T} V_{it} = V_{it} - \bar{V}_i$
- $\tilde{Y}_{it} = Y_{it} - \bar{Y}_i$: Outcome minus averaged outcome
- $\tilde{X}_{it} = X_{it} - \bar{X}_i$: Covariate(s) minus averaged covariate(s)
- $\tilde{D}_{it} = D_{it} - \bar{D}_i$: Treatment minus averaged treatment
- $\tilde{\alpha}_i = \alpha_i - \bar{\alpha}_i = 0$: Individual effect minus averaged individual effect
- $\tilde{\nu}_{it} = \nu_{it} - \bar{\nu}_i$: Error minus averaged error

We can then express the "de-meaned" version of our model:

$$\tilde{Y}_{it} = \tilde{X}_{it}\beta + \tau\tilde{D}_{it} + \tilde{\nu}_{it}$$

$\rightarrow$ This gives you the same result as "controlling for" α_i

We can extend these estimators to a more complex model

Consider the regression model with both individual and time effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \delta_t + \nu_{it}$$

We can extend these estimators to a more complex model

Consider the regression model with both individual and time effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \delta_t + \nu_{it}$$

The dummy variable method for controlling for α_i and δ_t is:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \sum_{i=1}^{N} \mathbf{1}[unit = i]_i + \sum_{t=1}^{T} \mathbf{1}[time = t]_t + \nu_{it}$$

We can extend these estimators to a more complex model

Consider the regression model with both individual and time effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \delta_t + \nu_{it}$$

The dummy variable method for controlling for α_i and δ_t is:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \sum_{i=1}^{N} \mathbf{1}[unit = i]_i + \sum_{t=1}^{T} \mathbf{1}[time = t]_t + \nu_{it}$$

De-meaning is more annoying. Now we de-mean by both time and unit:

Define

$$\tilde{\tilde{V}}_{it} = V_{it} - \frac{1}{T}\sum_{t=1}^{T} V_{it} - \frac{1}{N}\sum_{i=1}^{N} V_{it} + \frac{1}{N \times T}\sum_{i=1}^{N}\sum_{t=1}^{T} V_{it}$$

$$= V_{it} - \bar{V}_i - \bar{V}_t + \bar{\bar{V}}_{it}$$

We can extend these estimators to a more complex model

Consider the regression model with both individual and time effects:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \alpha_i + \delta_t + \nu_{it}$$

The dummy variable method for controlling for α_i and δ_t is:

$$Y_{it} = X_{it}\beta + \tau D_{it} + \sum_{i=1}^{N} \mathbf{1}[unit = i]_i + \sum_{t=1}^{T} \mathbf{1}[time = t]_t + \nu_{it}$$

De-meaning is more annoying. Now we de-mean by both time and unit:

Define

$$\tilde{\tilde{V}}_{it} = V_{it} - \frac{1}{T}\sum_{t=1}^{T} V_{it} - \frac{1}{N}\sum_{i=1}^{N} V_{it} + \frac{1}{N \times T}\sum_{i=1}^{N}\sum_{t=1}^{T} V_{it}$$

$$= V_{it} - \bar{V}_i - \bar{V}_t + \bar{\bar{V}}_{it}$$

The model is now:

$$\tilde{\tilde{Y}}_{it} = \tilde{\tilde{X}}_{it}\beta + \tau \tilde{\tilde{D}}_{it} + \tilde{\tilde{\nu}}_{it}$$

Connecting FE to DD

Remember that we started with a simple DD model:

$$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$$

Connecting FE to DD

Remember that we started with a simple DD model:

$$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$$

We can represent this as a regression as well:

$$Y_{it} = \alpha + \tau(Treat \times Post)_{it} + \gamma Treat_i + \delta Post_t + \beta X_{it} + \varepsilon_{it}$$

Connecting FE to DD

Remember that we started with a simple DD model:

$$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$$

We can represent this as a regression as well:

$$Y_{it} = \alpha + \tau(Treat \times Post)_{it} + \gamma Treat_i + \delta Post_t + \beta X_{it} + \varepsilon_{it}$$

In slightly different notation:

$$Y_{it} = \alpha + \tau D_{it} + \gamma Treat_i + \delta Post_t + \beta X_{it} + \varepsilon_{it}$$

Connecting FE to DD

Remember that we started with a simple DD model:

$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$

We can represent this as a regression as well:

$$Y_{it} = \alpha + \tau (Treat \times Post)_{it} + \gamma Treat_i + \delta Post_t + \beta X_{it} + \varepsilon_{it}$$

In slightly different notation:

$$Y_{it} = \alpha + \tau D_{it} + \gamma Treat_i + \delta Post_t + \beta X_{it} + \varepsilon_{it}$$

Now let's generalize:

$$Y_{it} = \tau D_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- α_i is a (set of) individual fixed effects, which captures $Treat_i$
- δ_t is a (set of) time fixed effects, which captures $Post_t$

Connecting FE to DD

Remember that we started with a simple DD model:

$\hat{\tau}^{DD} = (\bar{Y}(treat, post) - \bar{Y}(treat, pre)) - (\bar{Y}(untreat, post) - \bar{Y}(untreat, pre))$

We can represent this as a regression as well:

$$Y_{it} = \alpha + \tau(Treat \times Post)_{it} + \gamma Treat_i + \delta Post_t + \beta X_{it} + \varepsilon_{it}$$

In slightly different notation:

$$Y_{it} = \alpha + \tau D_{it} + \gamma Treat_i + \delta Post_t + \beta X_{it} + \varepsilon_{it}$$

Now let's generalize:

$$Y_{it} = \tau D_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- α_i is a (set of) individual fixed effects, which captures $Treat_i$
- δ_t is a (set of) time fixed effects, which captures $Post_t$

→ Note that we've lost α: this is now collinear with α_i

→ With just two units, α_i is just $\mathbf{1}[unit = i]$ and $\mathbf{1}[unit = j]$

→ With just two time periods, α_i is just $\mathbf{1}[time = 0]$ and $\mathbf{1}[time = 1]$

Extending DD to multiple treatment times

What happens if we have treated units who get treated at different times?

- Our new general formulation works!

$$Y_{it} = \tau D_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

Extending DD to multiple treatment times

What happens if we have treated units who get treated at different times?

- Our new general formulation works!

$$Y_{it} = \tau D_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- D_{it} can turn from 0 to 1 at different times for different units
- (Or can stay 0 always for some units)
- (Or can stay 1 always for some units)

Extending DD to multiple treatment times

What happens if we have treated units who get treated at different times?

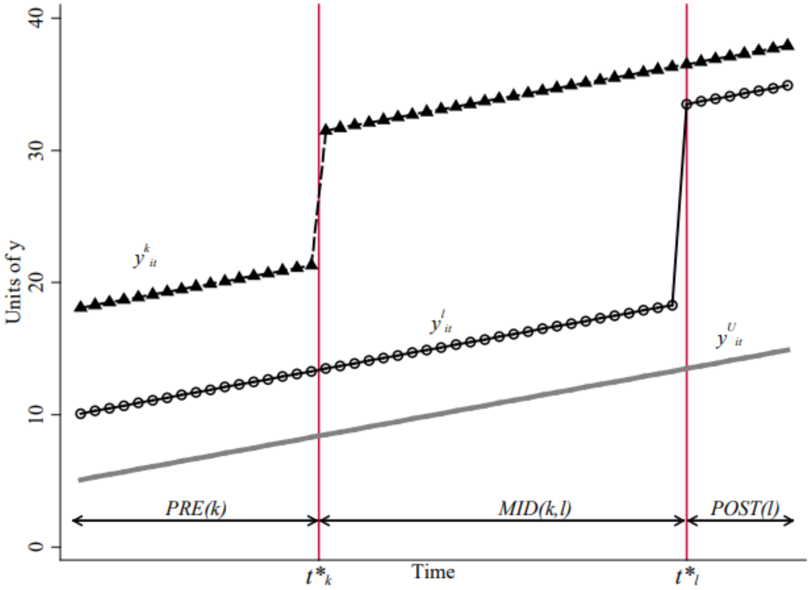
- Our new general formulation works!

$$Y_{it} = \tau D_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

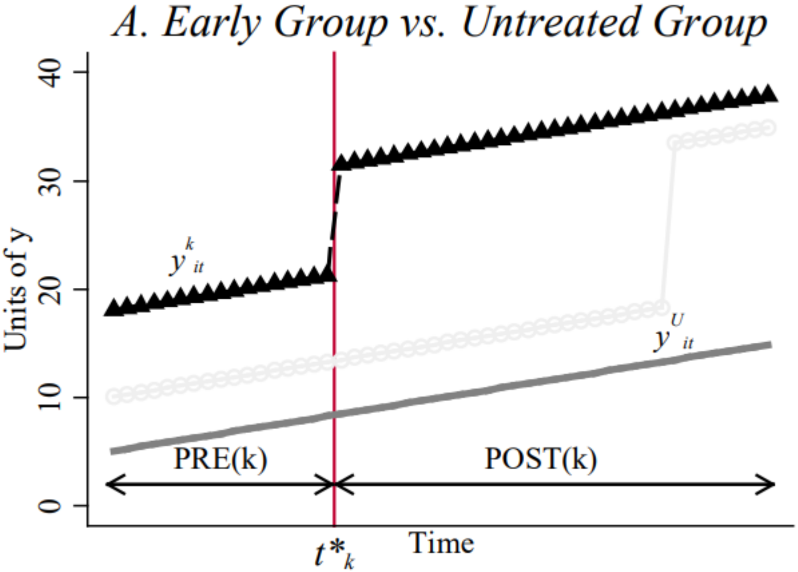
- D_{it} can turn from 0 to 1 at different times for different units
- (Or can stay 0 always for some units)
- (Or can stay 1 always for some units)

Important note: Running this specification gets you a **weighted average** of several comparisons. This may not be exactly what you want!

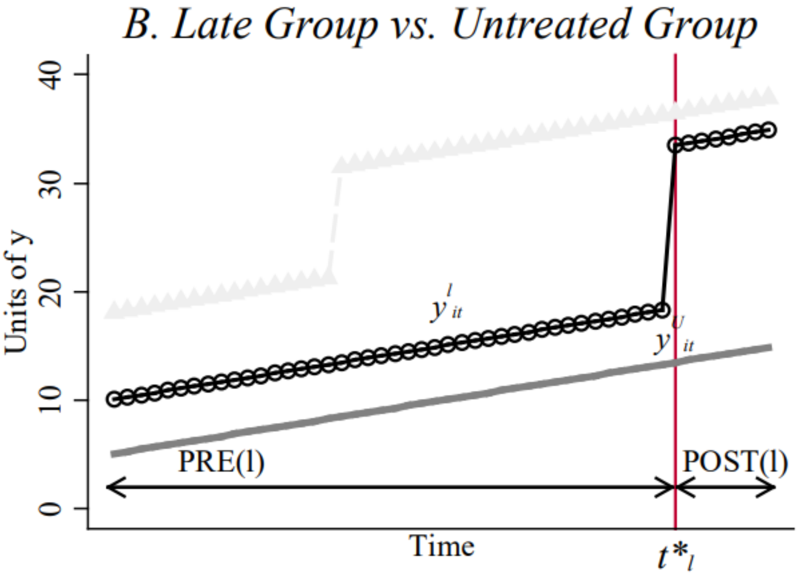
What does multiple-treatment-timing FE get us?



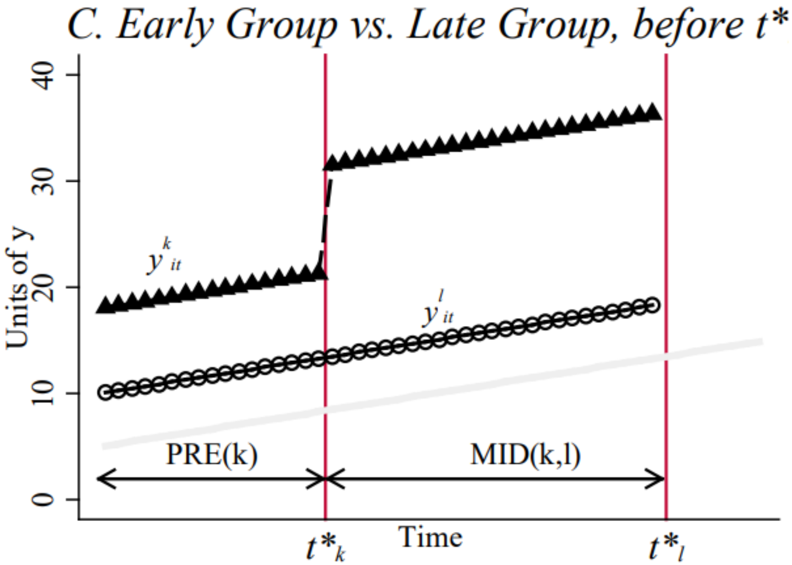
What does multiple-treatment-timing FE get us?



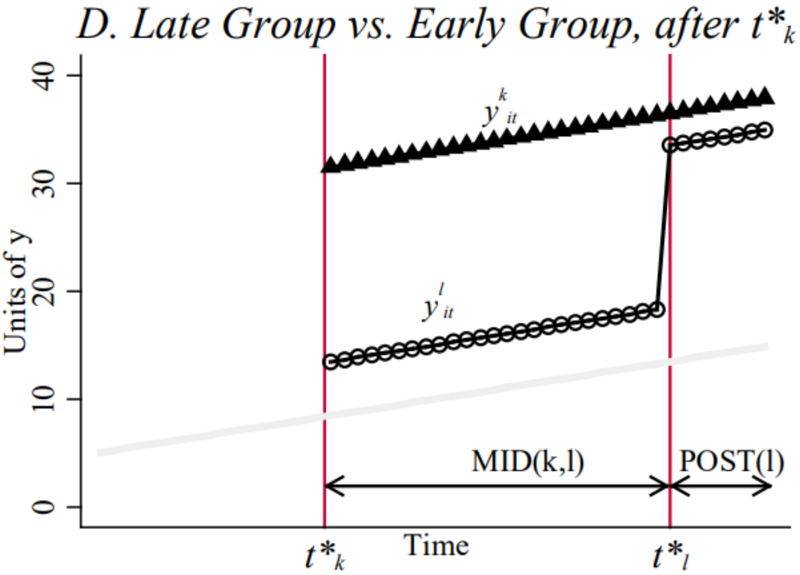
What does multiple-treatment-timing FE get us?



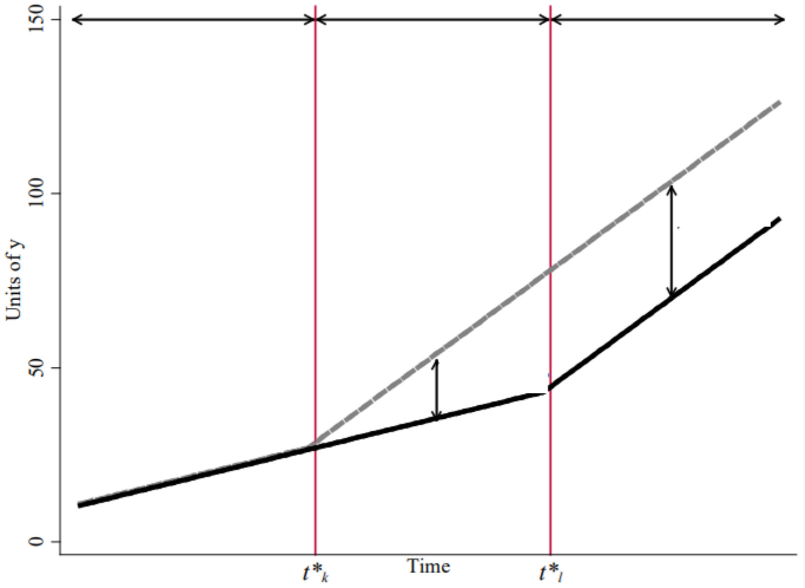
What does multiple-treatment-timing FE get us?



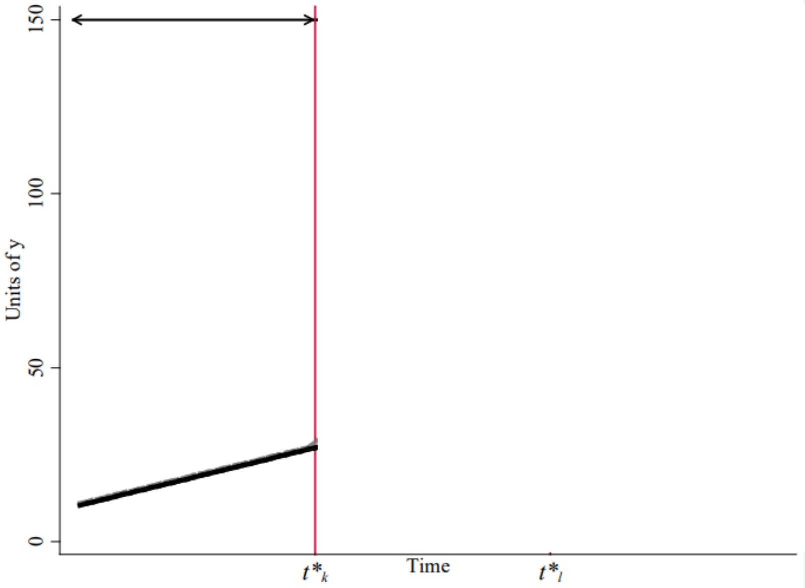
What does multiple-treatment-timing FE get us?



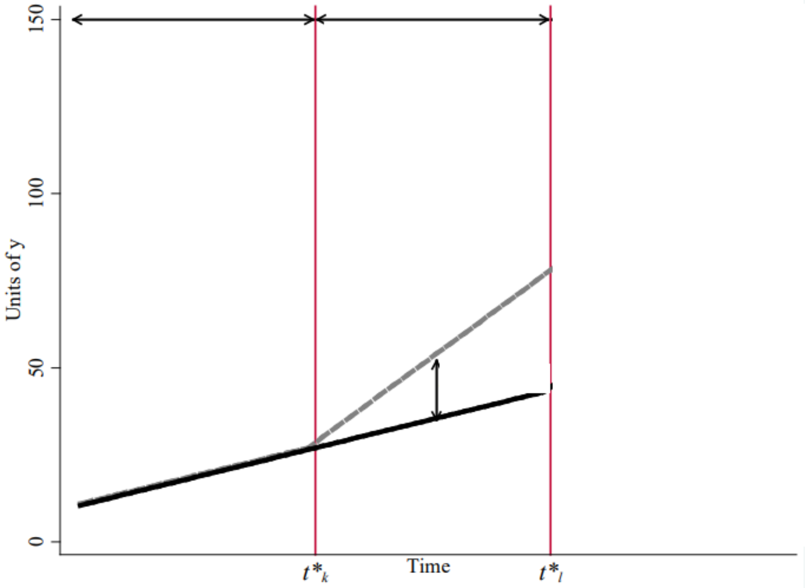
What does multiple-treatment-timing FE get us?



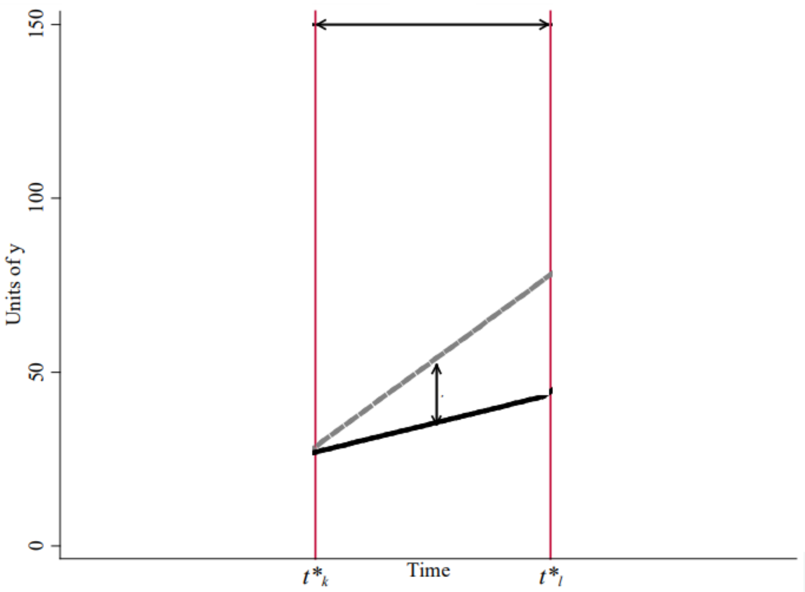
What does multiple-treatment-timing FE get us?



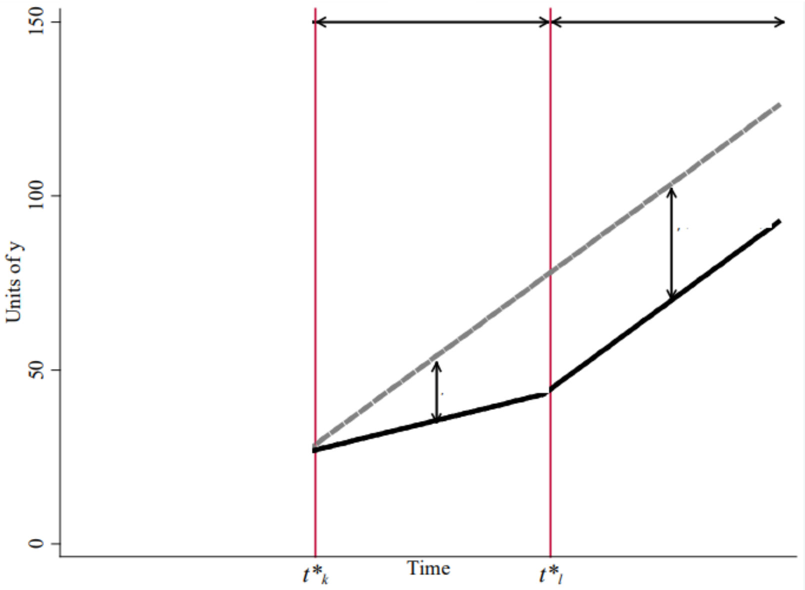
What does multiple-treatment-timing FE get us?



What does multiple-treatment-timing FE get us?



What does multiple-treatment-timing FE get us?



How do we do this right?

There are two main solutions to this problem:

1. Weighted balance test to make sure this isn't a problem
   - Beyond the scope of this class
2. Artificially "treat all units at the same time"

The event study design

An event study is a more general FE design:

Our standard FE regression model:

$$Y_{it} = \tau D_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- This imposes the constraint that $\tau_t = \tau$ for all t
- (And all i, but that's a problem for a different day)

The event study design

An event study is a more general FE design:

Our standard FE regression model:

$$Y_{it} = \tau D_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- This imposes the constraint that $\tau_t = \tau$ for all t
- (And all i, but that's a problem for a different day)

A more general model will allow differential effects over time:

$$Y_{it} = \sum_{r=0}^{R} \tau_r D_i \times \mathbf{1}[\text{periods post treatment} = r]_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

$\mathbf{1}[\text{periods post treatment} = r]_{it} = 1$ when we are r periods after treatment, 0 otherwise

The event study design

An event study is a more general FE design:

Our standard FE regression model:

$$Y_{it} = \tau D_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- This imposes the constraint that $\tau_t = \tau$ for all t
- (And all i, but that's a problem for a different day)

A more general model will allow differential effects over time:

$$Y_{it} = \sum_{r=0}^{R} \tau_r D_i \times \mathbf{1}[\text{periods post treatment} = r]_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

$\mathbf{1}[\text{periods post treatment} = r]_{it} = 1$ when we are r periods after treatment, 0 otherwise

- The τ_rs pick up the average treatment effect r periods after treatment

The event study design

In the general version of this model, we include pre-treatment "effects":

$$Y_{it} = \sum_{r=-S}^{R} \tau_r D_i \times \mathbf{1}[\text{periods to treatment} = r]_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- (Note the number of pre-treatment periods, S, and the number of post-treatment periods, R, need not be the same)

The event study design

In the general version of this model, we include pre-treatment "effects":

$$Y_{it} = \sum_{r=-S}^{R} \tau_r D_i \times \mathbf{1}[\text{periods to treatment} = r]_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- (Note the number of pre-treatment periods, S, and the number of post-treatment periods, R, need not be the same)

What's so great about this design?

- This "lines up" treatment at the same time for everyone
- We can still use fixed effects to soak up confounders
- We get a (partial) test of the identifying assumption

The event study design

In the general version of this model, we include pre-treatment "effects":

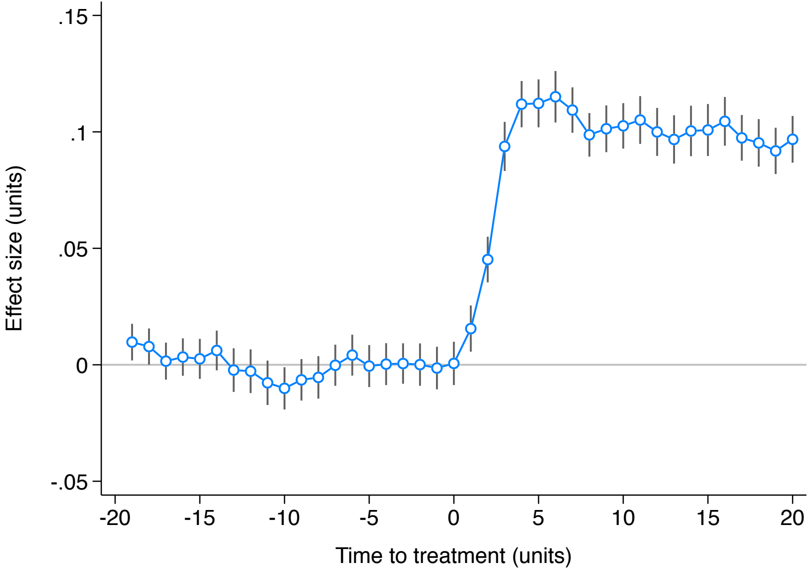
$$Y_{it} = \sum_{r=-S}^{R} \tau_r D_i \times \mathbf{1}[\text{periods to treatment} = r]_{it} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

- (Note the number of pre-treatment periods, S, and the number of post-treatment periods, R, need not be the same)

What's so great about this design?

- This "lines up" treatment at the same time for everyone
- We can still use fixed effects to soak up confounders
- We get a (partial) test of the identifying assumption
  - $\rightarrow$ We want pre-treatment τ_s's to be centered on 0 and not trending

Event studies should always be shown graphically



Cumulative effects

We're often interested in summing up effects over time:

- This is a simple extension of per-period event study effects
- And very powerful!
- We can use this to understand if treatment moves things in time
- Or happens and fades away
- Or remains important over time

Cumulative effects

We estimate cumulative effects with a distributed lag model:

$$Y_{it} = \sum_{s=0}^{S} \tau_s D_{i,t-s} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

where $D_{i,t-s}$ is an indicator equal to the treatment status in period $t - s$

Cumulative effects

We estimate cumulative effects with a distributed lag model:

$$Y_{it} = \sum_{s=0}^{S} \tau_s D_{i,t-s} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

where $D_{i,t-s}$ is an indicator equal to the treatment status in period $t - s$

- τ_0 gives the effect in the treatment period
- τ_1 gives the marginal effect of treatment 1 period later
- τ_2 gives the marginal effect of treatment 2 periods later

  ...
- τ_S gives the marginal effect of treatment S periods later

Cumulative effects

We estimate cumulative effects with a distributed lag model:

$$Y_{it} = \sum_{s=0}^{S} \tau_s D_{i,t-s} + \alpha_i + \delta_t + \beta X_{it} + \varepsilon_{it}$$

where $D_{i,t-s}$ is an indicator equal to the treatment status in period $t-s$

- τ_0 gives the effect in the treatment period
- τ_1 gives the marginal effect of treatment 1 period later
- τ_2 gives the marginal effect of treatment 2 periods later

  ...

- τ_S gives the marginal effect of treatment S periods later

→ The cumulative effect q periods after treatment is:

$$T_q = \sum_{s=0}^{q} \tau_s$$

Recap

TL;DR:

1. We like the difference-in-differences approach a lot
2. We discussed estimation with fixed effects
3. And covered the event study version