Econometrics III: Applied Methods — NovaSBE, 2026

Problem Set 3 Solutions: Difference-in-Differences

Exercise A: DiD Foundations

1.1 *(1 pt)* Compute the DiD estimate. Write out the parallel trends assumption for this setting and give one concrete reason why it might fail.

Solution:

DiD Estimate:

$$\begin{aligned}\hat{\tau}^{DiD} &= (\bar{Y}_{NJ,after} - \bar{Y}_{NJ,before}) - (\bar{Y}_{PA,after} - \bar{Y}_{PA,before}) \\ &= (21.03 - 20.44) - (21.17 - 23.33) \\ &= 0.59 - (-2.16) \\ &= \boxed{2.75 \text{ FTE per restaurant}}\end{aligned}$$

The estimate suggests the minimum wage increase *raised* employment by about 2.75 full-time equivalent workers per restaurant, contrary to the standard competitive model prediction.

Parallel Trends Assumption:

$$\mathbb{E}[Y_{NJ,post}(0) - Y_{NJ,pre}(0)] = \mathbb{E}[Y_{PA,post}(0) - Y_{PA,pre}(0)]$$

In words: absent the minimum wage increase, employment in New Jersey would have evolved on the same trajectory as employment in Pennsylvania.

Why it might fail: New Jersey and Pennsylvania may have different underlying economic trends. For example:

- NJ's economy was growing faster than PA's in 1992 (stronger consumer demand)
- Differential regional shocks (e.g., opening/closing of major employers)
- Different seasonal patterns in fast-food employment across states
- Pre-existing differences in fast-food market saturation

1.2 *(1.5 pts)* Suppose employment in both states follows the model:

$$Y_{st} = \alpha_s + \delta_t + \beta \cdot D_{st} + \gamma_s \cdot t + \varepsilon_{st}$$

where γ_s is a state-specific linear trend. Show that if $\gamma_{NJ} \neq \gamma_{PA}$, the DiD estimator is biased for β. Derive the bias term, and explain it.

Solution:

Under this model, the expected outcomes are (ignoring ε):

$$\begin{aligned}\mathbb{E}[Y_{NJ,post}] &= \alpha_{NJ} + \delta_{post} + \beta + \gamma_{NJ} \cdot t_{post} \\ \mathbb{E}[Y_{NJ,pre}] &= \alpha_{NJ} + \delta_{pre} + \gamma_{NJ} \cdot t_{pre} \\ \mathbb{E}[Y_{PA,post}] &= \alpha_{PA} + \delta_{post} + \gamma_{PA} \cdot t_{post} \\ \mathbb{E}[Y_{PA,pre}] &= \alpha_{PA} + \delta_{pre} + \gamma_{PA} \cdot t_{pre}\end{aligned}$$

The DiD estimator computes:

$$\hat{\tau}^{DiD} = (\bar{Y}_{NJ,post} - \bar{Y}_{NJ,pre}) - (\bar{Y}_{PA,post} - \bar{Y}_{PA,pre})$$

Taking expectations:

$$\begin{aligned}\mathbb{E}[\hat{\tau}^{DiD}] &= [(\delta_{post} - \delta_{pre}) + \beta + \gamma_{NJ}(t_{post} - t_{pre})] - [(\delta_{post} - \delta_{pre}) + \gamma_{PA}(t_{post} - t_{pre})] \\ &= \beta + (\gamma_{NJ} - \gamma_{PA})(t_{post} - t_{pre})\end{aligned}$$

Therefore:

$$\boxed{\text{Bias} = (\gamma_{NJ} - \gamma_{PA}) \cdot \Delta t}$$

where $\Delta t = t_{post} - t_{pre}$.

Explanation: The bias equals the *differential trend* between states multiplied by the time elapsed. If NJ has a stronger underlying growth trend ($\gamma_{NJ} > \gamma_{PA}$), DiD attributes some of this natural growth to the treatment effect, biasing the estimate upward.

> **1.3** *(0.5 pts)* A skeptic argues: "New Jersey's economy was stronger than Pennsylvania's in 1992, so employment would have grown faster in NJ even without the minimum wage increase." Translate this concern into the model from 1.2. What sign would $\gamma_{NJ} - \gamma_{PA}$ have? Would DiD overestimate or underestimate the employment effect of the minimum wage?

Solution:

The skeptic's concern translates to:

$$\gamma_{NJ} > \gamma_{PA} \quad \Rightarrow \quad \gamma_{NJ} - \gamma_{PA} > 0$$

From 1.2, the bias is $(\gamma_{NJ} - \gamma_{PA}) \cdot \Delta t > 0$.

Therefore, DiD would **overestimate** the employment effect of the minimum wage.

Interpretation: If NJ employment was naturally growing faster than PA, some of that underlying growth would be incorrectly attributed to the minimum wage policy. The true causal effect of the minimum wage would be smaller (perhaps even negative) than the DiD estimate of 2.75 suggests.

Note: This is exactly why parallel trends is such a critical assumption. The DiD estimate conflates the causal effect with any differential trends that would have occurred anyway.

> **1.4** *(1 pt)* Suppose workers can commute across the NJ–PA border for work. After NJ raises its minimum wage, some PA residents may seek jobs in NJ (where wages are now higher), while some NJ firms may relocate to PA (to avoid the higher wage). How might this cross-border mobility affect the DiD estimate? Would it lead to overestimation or underestimation of the employment effect?

Solution:

Cross-border mobility creates **spillover effects** that contaminate the control group.

Mechanisms:

1. **Labor supply shift to NJ:** PA workers seeking higher-wage NJ jobs increases labor supply in NJ. This could:
   - Make it easier for NJ firms to fill positions (offsetting job losses)

- Push down effective wages toward the new minimum

2. **Firm relocation to PA:** NJ firms moving to PA increases PA employment. This:
    - Reduces NJ employment (captured in treatment effect)
    - Increases PA employment (contaminates control)

Effect on DiD estimate:
The firm relocation effect is key. If some employment "moves" from NJ to PA:

- NJ employment falls more than it would without mobility
- PA employment rises more than it would without mobility

The DiD estimate is:

$$\hat{\tau}^{DiD} = \underbrace{(\Delta Y_{NJ})}_{\text{includes loss to PA}} - \underbrace{(\Delta Y_{PA})}_{\text{includes gain from NJ}}$$

Both effects make the DiD estimate **more negative** (or less positive) than the true effect would be in a closed economy.

Conclusion: Cross-border mobility leads to **underestimation** of any positive employment effect (or overestimation of negative effects). The DiD captures relocation across borders, not just the direct employment effect within NJ.

This violates SUTVA—the control group (PA) is affected by treatment in NJ.

Exercise B: Formal Analysis of Threats to Identification

B.1 Ashenfelter's Dip and Mean Reversion

1.1 *(1 pt)* Explain intuitively why participants might enroll if $Y_{i1}(0) < c$ in this kind of context. Explain intuitively why $\mathbb{E}[\eta_{i1}|D_i = 1] < 0$.

Solution:

Part 1: Why workers enroll when $Y_{i1}(0) < c$

The selection rule $D_i = \mathbb{1}(Y_{i1}(0) < c)$ reflects a plausible behavioral model:

- Workers with low current earnings have more to gain from training (higher returns)
- Low earnings may signal a need for skill upgrading
- Workers "hitting bottom" may be more motivated to invest in human capital
- Government programs often target workers with earnings below some threshold

In essence, workers experiencing hard times are more likely to seek help through training programs.

Part 2: Why $\mathbb{E}[\eta_{i1}|D_i = 1] < 0$

Workers enroll if $Y_{i1}(0) < c$, which means:

$$\mu_i + \delta_1 + \eta_{i1} < c$$

Since η_{i1} is mean-zero in the population, workers with *unusually low* realizations of η_{i1} are more likely to satisfy this condition and enroll. Selection is based partly on a negative transitory shock.

Intuition: Workers who had a "bad year" (negative η_{i1})—perhaps due to temporary illness, family emergency, or bad luck in the job market—are more likely to have low earnings and thus more likely to seek training.

Therefore:

$$\mathbb{E}[\eta_{i1}|D_i = 1] < 0$$

Implication: Trainees are not a random sample; they are selected partly based on having an unusually bad transitory shock in period 1.

1.2 *(1.5 pts)* Assume the true treatment effect is τ (so $Y_{i2} = Y_{i2}(0) + \tau \cdot D_i$). Derive an expression for the DiD estimator, defined as:

$$\hat{\tau}^{DiD} = \mathbb{E}[Y_{i2} - Y_{i1}|D_i = 1] - \mathbb{E}[Y_{i2} - Y_{i1}|D_i = 0]$$

Show that DiD is biased and determine the sign of the bias.

Solution:

Step 1: Derive the change in potential outcomes for each group.

Using $Y_{it}(0) = \mu_i + \delta_t + \eta_{it}$:

For trainees ($D_i = 1$):

$$\begin{aligned}\mathbb{E}[Y_{i2}(0) - Y_{i1}(0)|D_i = 1] &= \mathbb{E}[\mu_i + \delta_2 + \eta_{i2} - \mu_i - \delta_1 - \eta_{i1}|D_i = 1] \\ &= \delta_2 - \delta_1 + \mathbb{E}[\eta_{i2}|D_i = 1] - \mathbb{E}[\eta_{i1}|D_i = 1]\end{aligned}$$

Since η_{i2} is independent of both η_{i1} (serially uncorrelated) and μ_i (by assumption), and selection D_i depends only on (μ_i, η_{i1}):

$$\mathbb{E}[\eta_{i2}|D_i = 1] = \mathbb{E}[\eta_{i2}] = 0$$

Therefore:

$$\mathbb{E}[Y_{i2}(0) - Y_{i1}(0)|D_i = 1] = \delta_2 - \delta_1 - \mathbb{E}[\eta_{i1}|D_i = 1]$$

For non-trainees ($D_i = 0$): By identical logic:

$$\mathbb{E}[Y_{i2}(0) - Y_{i1}(0)|D_i = 0] = \delta_2 - \delta_1 - \mathbb{E}[\eta_{i1}|D_i = 0]$$

Step 2: Derive the DiD estimator.
The DiD estimator compares observed outcomes. Since $Y_{i2} = Y_{i2}(0) + \tau D_i$:

$$\begin{aligned}
\hat{\tau}^{DiD} &= \mathbb{E}[Y_{i2} - Y_{i1}|D_i = 1] - \mathbb{E}[Y_{i2} - Y_{i1}|D_i = 0] \\
&= \mathbb{E}[Y_{i2}(0) + \tau - Y_{i1}(0)|D_i = 1] - \mathbb{E}[Y_{i2}(0) - Y_{i1}(0)|D_i = 0] \\
&= \tau + \mathbb{E}[Y_{i2}(0) - Y_{i1}(0)|D_i = 1] - \mathbb{E}[Y_{i2}(0) - Y_{i1}(0)|D_i = 0]
\end{aligned}$$

Substituting from Step 1:

$$\begin{aligned}
\hat{\tau}^{DiD} &= \tau + (\delta_2 - \delta_1 - \mathbb{E}[\eta_{i1}|D_i = 1]) - (\delta_2 - \delta_1 - \mathbb{E}[\eta_{i1}|D_i = 0]) \\
&= \tau - (\mathbb{E}[\eta_{i1}|D_i = 1] - \mathbb{E}[\eta_{i1}|D_i = 0])
\end{aligned}$$

Define $\Delta\eta = \mathbb{E}[\eta_{i1}|D_i = 1] - \mathbb{E}[\eta_{i1}|D_i = 0]$. Then:

$$\boxed{\hat{\tau}^{DiD} = \tau - \Delta\eta}$$

Step 3: Sign of the bias.
From Q1.1, $\mathbb{E}[\eta_{i1}|D_i = 1] < 0$ (trainees had negative transitory shocks). Non-trainees on average have $\mathbb{E}[\eta_{i1}|D_i = 0] > 0$ (they didn't need training because they were doing okay). Therefore:

$$\Delta\eta = \mathbb{E}[\eta_{i1}|D_i = 1] - \mathbb{E}[\eta_{i1}|D_i = 0] < 0$$

So Bias $= -\Delta\eta > 0$. **DiD overestimates** the true training effect.

1.3 *(0.5 pts)* This bias is sometimes called "mean reversion bias." Explain why: what happens to the transitory component η_{it} between periods 1 and 2, and why does this look like a treatment effect?

Solution:

Mean reversion: Since η_{it} is transitory with $\text{Cov}(\eta_{i1}, \eta_{i2}) = 0$, workers who had unusually low η_{i1} will, on average, have η_{i2} closer to zero. Their earnings naturally "revert to the mean."

For trainees:

- Period 1: $\mathbb{E}[\eta_{i1}|D_i = 1] < 0$ (unusually bad)
- Period 2: $\mathbb{E}[\eta_{i2}|D_i = 1] = 0$ (back to normal)

Their earnings increase from period 1 to 2 by an amount $-\mathbb{E}[\eta_{i1}|D_i = 1] > 0$ even *without any treatment effect.*

Why it looks like a treatment effect: We observe trainees' earnings rising after training. But part of this increase is just their transitory luck returning to normal—not the causal effect of training. DiD conflates this mechanical rebound with the true program impact.

Graphically: Trainees start from a "dip" (Ashenfelter's dip) below their long-run average. Even without training, they would recover. DiD incorrectly attributes this recovery to the program.

1.4 *(1 pt)* Would this bias show up in a pre-trends test if we had data from period 0?

Solution:

Yes, the pre-trends test would detect the problem.

Since $D_i = \mathbb{1}(\mu_i + \delta_1 + \eta_{i1} < c)$, selection depends only on period-1 shocks. The period-0 shock η_{i0} is independent of D_i.

The pre-trend is:

$$\begin{aligned}
&\mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|D_i = 1] - \mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|D_i = 0] \\
&= (\delta_1 - \delta_0 + \mathbb{E}[\eta_{i1}|D_i = 1] - \underbrace{\mathbb{E}[\eta_{i0}|D_i = 1]}_{=0}) - (\delta_1 - \delta_0 + \mathbb{E}[\eta_{i1}|D_i = 0] - \underbrace{\mathbb{E}[\eta_{i0}|D_i = 0]}_{=0}) \\
&= \mathbb{E}[\eta_{i1}|D_i = 1] - \mathbb{E}[\eta_{i1}|D_i = 0] = \Delta\eta \neq 0
\end{aligned}$$

The differential pre-trend equals $\Delta\eta < 0$, showing trainees' earnings *falling* relative to non-trainees just before treatment. This is the visible signature of Ashenfelter's dip—a warning sign that selection is occurring based on transitory shocks.

B.2 Anticipation Effects

2.1 *(1 pt)* Show that the researcher's DiD estimate equals $\tau^{post} - \tau^{antic}$, not τ^{post}.

Solution:

The true potential outcome model is:

$$Y_{it}(d) = \alpha_i + \delta_t + \tau^{post} \cdot d \cdot \mathbb{1}(t \geq 1) + \tau^{antic} \cdot d \cdot \mathbb{1}(t = 0) + \varepsilon_{it}$$

The researcher, unaware of anticipation, defines $t = 0$ as "pre" and $t = 1$ as "post."

For treated units ($D_i = 1$):

$$\begin{aligned}
\mathbb{E}[Y_{i,1}|D_i = 1] &= \alpha + \delta_1 + \tau^{post} \\
\mathbb{E}[Y_{i,0}|D_i = 1] &= \alpha + \delta_0 + \tau^{antic}
\end{aligned}$$

For control units ($D_i = 0$):

$$\begin{aligned}
\mathbb{E}[Y_{i,1}|D_i = 0] &= \alpha + \delta_1 \\
\mathbb{E}[Y_{i,0}|D_i = 0] &= \alpha + \delta_0
\end{aligned}$$

The DiD estimate:

$$\begin{aligned}
\hat{\tau}^{DiD} &= (\bar{Y}_{1,1} - \bar{Y}_{1,0}) - (\bar{Y}_{0,1} - \bar{Y}_{0,0}) \\
&= [(\delta_1 + \tau^{post}) - (\delta_0 + \tau^{antic})] - [(\delta_1) - (\delta_0)] \\
&= \tau^{post} - \tau^{antic}
\end{aligned}$$

$$\boxed{\hat{\tau}^{DiD} = \tau^{post} - \tau^{antic}}$$

The estimate captures the *change* in the treatment effect from the anticipation period to the post period, not the level of the post-treatment effect.

2.2 *(1 pt)* The researcher runs an event study with periods $t \in \{-1, 0, 1, 2\}$, using $t = -1$ as the reference period. Derive the population values of β_0, β_1, and β_2.

Solution:

The event study specification estimates:

$$Y_{it} = \alpha_i + \delta_t + \sum_{k \neq -1} \beta_k \cdot \mathbb{1}(t = k) \cdot D_i + \varepsilon_{it}$$

The coefficient β_k captures the differential outcome for treated vs. control at time k, relative to $k = -1$.

Under the true model with anticipation at $t = 0$ and treatment at $t \geq 1$:

β_0**: Differential change from** $t = -1$ **to** $t = 0$

$$\begin{aligned}\beta_0 &= \mathbb{E}[Y_{i,0} - Y_{i,-1} | D_i = 1] - \mathbb{E}[Y_{i,0} - Y_{i,-1} | D_i = 0] \\ &= [(0 + \tau^{antic}) - 0] - [0 - 0] = \boxed{\tau^{antic}}\end{aligned}$$

β_1**: Differential change from** $t = -1$ **to** $t = 1$

$$\begin{aligned}\beta_1 &= \mathbb{E}[Y_{i,1} - Y_{i,-1} | D_i = 1] - \mathbb{E}[Y_{i,1} - Y_{i,-1} | D_i = 0] \\ &= \tau^{post} - 0 = \boxed{\tau^{post}}\end{aligned}$$

β_2**: Differential change from** $t = -1$ **to** $t = 2$

$$\begin{aligned}\beta_2 &= \mathbb{E}[Y_{i,2} - Y_{i,-1} | D_i = 1] - \mathbb{E}[Y_{i,2} - Y_{i,-1} | D_i = 0] \\ &= \tau^{post} - 0 = \boxed{\tau^{post}}\end{aligned}$$

(Assuming the post-treatment effect is constant at τ^{post} for $t \geq 1$.)

2.3 *(0.5 pts)* Looking at the event study, how would the researcher detect anticipation effects? What pattern would signal this problem?

Solution:

Detection: The researcher would see $\beta_0 \neq 0$ even though $t = 0$ is classified as "pre-treatment."

Pattern signaling anticipation:

- Pre-treatment coefficients $(\beta_{-2}, \beta_{-1}, \ldots)$ are all zero
- At $t = 0$: $\beta_0 \neq 0$ (a "jump" before official implementation)
- At $t \geq 1$: $\beta_1, \beta_2, \ldots$ at a different level than β_0

Graphically: The event study would show a non-zero coefficient at $t = 0$ (the announcement period), distinct from both:

1. The flat zero line in earlier pre-periods
2. The post-treatment effect in later periods

Recommendation: If anticipation is suspected, the researcher should redefine $t = 0$ as part of the treatment period, or use $t = -2$ as the reference period and examine whether β_{-1} and β_0 are both non-zero.

2.4 *(1 pt)* Suppose we are studying a corporate tax increase. If firms accelerate investment to $t = 0$ to avoid higher taxes in $t \geq 1$, what are the signs of τ^{antic} and τ^{post}? How does this affect the DiD estimate?

Solution:

Signs:

- $\tau^{antic} > 0$: Investment *increases* at $t = 0$ as firms pull forward spending to avoid higher future taxes
- $\tau^{post} < 0$: Investment *decreases* at $t \geq 1$ due to (a) higher taxes discouraging investment and (b) depletion from the anticipatory surge

Effect on DiD:

$$\hat{\tau}^{DiD} = \tau^{post} - \tau^{antic} = (\text{negative}) - (\text{positive}) < 0$$

The DiD estimate is **biased downward** (more negative than τ^{post} alone).

Intuition: The researcher compares $t = 1$ to $t = 0$, but $t = 0$ is artificially inflated by anticipatory behavior. The measured drop from $t = 0$ to $t = 1$ reflects both:

1. The true negative effect of higher taxes on investment
2. The reversal of the anticipatory surge

DiD **overstates** how much the tax increase reduced investment because it measures from an inflated baseline.

B.3 Heterogeneous Trends and the Bias Formula

3.1 *(1 pt)* Show that parallel trends holds if and only if $\mathbb{E}[\gamma_i|D_i = 1] = \mathbb{E}[\gamma_i|D_i = 0]$.

Solution:

Under the model $Y_{it}(0) = \alpha_i + \delta_t + \gamma_i \cdot t + \varepsilon_{it}$, the change in untreated potential outcomes from t to t' is:

$$\begin{aligned}\mathbb{E}[Y_{it'}(0) - Y_{it}(0)|D_i = d] &= \mathbb{E}[\alpha_i + \delta_{t'} + \gamma_i t' - \alpha_i - \delta_t - \gamma_i t|D_i = d] \\ &= (\delta_{t'} - \delta_t) + \mathbb{E}[\gamma_i|D_i = d] \cdot (t' - t)\end{aligned}$$

Parallel trends requires:

$$\mathbb{E}[Y_{it'}(0) - Y_{it}(0)|D_i = 1] = \mathbb{E}[Y_{it'}(0) - Y_{it}(0)|D_i = 0]$$

Substituting:

$$(\delta_{t'} - \delta_t) + \bar{\gamma}_1(t' - t) = (\delta_{t'} - \delta_t) + \bar{\gamma}_0(t' - t)$$

This simplifies to:

$$\bar{\gamma}_1(t' - t) = \bar{\gamma}_0(t' - t)$$

For any $t' \neq t$:

$$\boxed{\bar{\gamma}_1 = \bar{\gamma}_0}$$

where $\bar{\gamma}_d = \mathbb{E}[\gamma_i | D_i = d]$.

3.2 *(1 pt)* Define $\bar{\gamma}_1 = \mathbb{E}[\gamma_i | D_i = 1]$ and $\bar{\gamma}_0 = \mathbb{E}[\gamma_i | D_i = 0]$. Show that for a 2×2 DiD comparing periods t and t' (with $t' > t \geq T^*$):

$$\text{Bias} = (\bar{\gamma}_1 - \bar{\gamma}_0) \cdot (t' - t)$$

Solution:

The 2×2 DiD estimator compares treated and control across two post-treatment periods:

$$\hat{\tau}^{DiD} = (\bar{Y}_{1,t'} - \bar{Y}_{1,t}) - (\bar{Y}_{0,t'} - \bar{Y}_{0,t})$$

For treated units (with true effect τ, assumed constant):

$$\begin{aligned}\mathbb{E}[\bar{Y}_{1,t'} - \bar{Y}_{1,t}] &= (\delta_{t'} - \delta_t) + \bar{\gamma}_1(t' - t) + (\tau - \tau) \\ &= (\delta_{t'} - \delta_t) + \bar{\gamma}_1(t' - t)\end{aligned}$$

For control units:

$$\mathbb{E}[\bar{Y}_{0,t'} - \bar{Y}_{0,t}] = (\delta_{t'} - \delta_t) + \bar{\gamma}_0(t' - t)$$

Therefore:

$$\begin{aligned}\mathbb{E}[\hat{\tau}^{DiD}] &= [(\delta_{t'} - \delta_t) + \bar{\gamma}_1(t' - t)] - [(\delta_{t'} - \delta_t) + \bar{\gamma}_0(t' - t)] \\ &= (\bar{\gamma}_1 - \bar{\gamma}_0)(t' - t)\end{aligned}$$

If we're estimating the effect τ, the true parameter is 0 (comparing two post-periods), but we get:

$$\boxed{\text{Bias} = (\bar{\gamma}_1 - \bar{\gamma}_0)(t' - t)}$$

Note: If comparing pre ($t < T^*$) to post ($t' \geq T^*$), the DiD identifies $\tau + (\bar{\gamma}_1 - \bar{\gamma}_0)(t' - t)$.

3.3 *(0.5 pts)* Using the result from 3.2, explain why DiD bias grows with the length of the post-treatment window. What does this imply for research design?

Solution:

Bias grows with time: The bias formula $(\bar{\gamma}_1 - \bar{\gamma}_0)(t' - t)$ is proportional to $(t' - t)$, the length of the observation window.

Why:

- Differential trends accumulate over time
- A small annual difference in trends ($\bar{\gamma}_1 - \bar{\gamma}_0 = 0.5\%$/year) becomes large over many years
- Over 10 years: $0.5\% \times 10 = 5\%$ bias

Implications for research design:

1. **Shorter windows are safer:** Compare periods close together to minimize trend accumulation

2. **Long-run effects are harder to identify:** Claims about effects 5–10 years post-treatment require stronger assumptions

3. **Use event studies:** Examine whether treatment effects grow implausibly large over time (may indicate trend contamination)

4. **Pre-trends tests are more informative with longer pre-periods:** If $\bar{\gamma}_1 \neq \bar{\gamma}_0$, this should be visible in multiple pre-periods

5. **Consider trend controls:** If differential trends are linear and constant, controlling for group-specific linear trends may help (but see D.3 for caveats)

> **3.4** *(1 pt)* Suppose treatment is more likely in regions with stronger economic growth (higher γ_i). What is the sign of the bias? Relate this to omitted variable bias: what is the "omitted variable" and why is it correlated with treatment?

Solution:

If treatment targets high-growth regions:

$$\bar{\gamma}_1 > \bar{\gamma}_0 \quad \Rightarrow \quad \text{Bias} = (\bar{\gamma}_1 - \bar{\gamma}_0)(t' - t) > 0$$

DiD overestimates the treatment effect—some of the observed "effect" is just underlying growth that would have occurred anyway.

OVB interpretation:

- **Omitted variable:** The unit-specific trend γ_i (or equivalently, $\gamma_i \times t$)
- **Why correlated with treatment:** Policies are often targeted at or adopted by high-growth areas (e.g., cities that attract investment also adopt progressive policies)
- **Why correlated with outcome:** Higher γ_i directly increases Y_{it}

Standard OVB formula:

$$\text{Bias} = \underbrace{\gamma}_{\text{effect of trend on } Y} \times \underbrace{\delta}_{\text{diff. in trends by treatment}}$$

In our notation: $\gamma = (t' - t)$ (how much trends matter) and $\delta = (\bar{\gamma}_1 - \bar{\gamma}_0)$ (differential trends by treatment status).

Policy example: Enterprise zones are often placed in areas showing early signs of revitalization. DiD may find "effects" that are really just continuation of pre-existing improvement.

Exercise C: Pre-testing Pitfalls (Roth, 2022)

C.1 The Power Problem

1.1–1.4 *(2 pts)* Analysis of pre-trends testing power.

Solution:

Results from R analysis of `pretrends_power.csv`:

1.1 Fraction passing pre-trends test:

Testing $H_0 : \beta_{pre} = 0$ at the 5% level:

$$\text{Reject if } \left| \frac{\hat{\beta}_{pre}}{SE_{pre}} \right| > 1.96$$

Result: **82.2%** of studies "pass" (fail to reject the null of zero pre-trend).

This is striking because the true pre-trend is $0.5 \neq 0$—parallel trends is *violated* in every study!

1.2 Power against true pre-trend:

The fraction rejecting a pre-trend of magnitude ≥ 0.5 is only **17.8%**.

Power is extremely low—we fail to detect the violation 82% of the time.

1.3 Breakdown pre-trend:

The "breakdown pre-trend" is the largest violation consistent with our estimate:

$$\text{Breakdown} = |\hat{\beta}_{pre}| + 1.96 \times SE_{pre}$$

Average breakdown: $\approx$ **1.56**

This is *larger* than the true treatment effect of 1.0! A pre-trend violation exceeding the effect size cannot be ruled out.

1.4 Why non-significant pre-trend is weak evidence:

1. **Low power:** Standard pre-trends tests are designed to detect *differences from zero*, not to confirm parallel trends holds
2. **Noise swamps signal:** With typical sample sizes, we can only detect large, obvious violations
3. **Absence of evidence $\neq$ evidence of absence:** Not rejecting $H_0 : \beta_{pre} = 0$ doesn't mean β_{pre} is actually zero
4. **Relevant comparison:** What matters is whether the pre-trend is large *relative to the treatment effect*—not whether it's statistically different from zero

Key insight (Roth, 2022): If the confidence interval for the pre-trend includes violations that could explain a substantial fraction of the estimated treatment effect, the pre-trends test provides little reassurance.

C.2 Selection Bias from Pre-testing

2.1–2.3 *(2 pts)* Bias from conditioning on passing pre-trends.

Solution:

Continuing with `pretrends_power.csv`:

2.1 Average $\hat{\beta}_{pre}$ and $\hat{\beta}_{post}$ among studies that pass:

Among studies passing the pre-trends test:

- $\bar{\beta}_{pre} \approx \mathbf{0.32}$ (true pre-trend: 0.5)
- $\bar{\beta}_{post} \approx \mathbf{1.59}$ (true effect: $1.0 + 0.5 = 1.5$)

The pre-trend estimate among passers (0.32) is **biased downward** from the truth (0.5)—studies that pass have **negatively selected** pre-trend estimates. The post-treatment effect is biased upward by about 0.09 relative to the unconditional average (≈ 1.5), reflecting selection on correlated noise.

2.2 Intuition for the bias:

Conditioning on passing ($|\hat{\beta}_{pre}| < 1.96 \times SE$) selects studies where:

- The pre-trend estimate happened to be small (by chance)
- Noise pushed $\hat{\beta}_{pre}$ toward zero

Because $\hat{\beta}_{pre}$ and $\hat{\beta}_{post}$ share common sampling variation, selecting on $\hat{\beta}_{pre}$ distorts the distribution of $\hat{\beta}_{post}$.

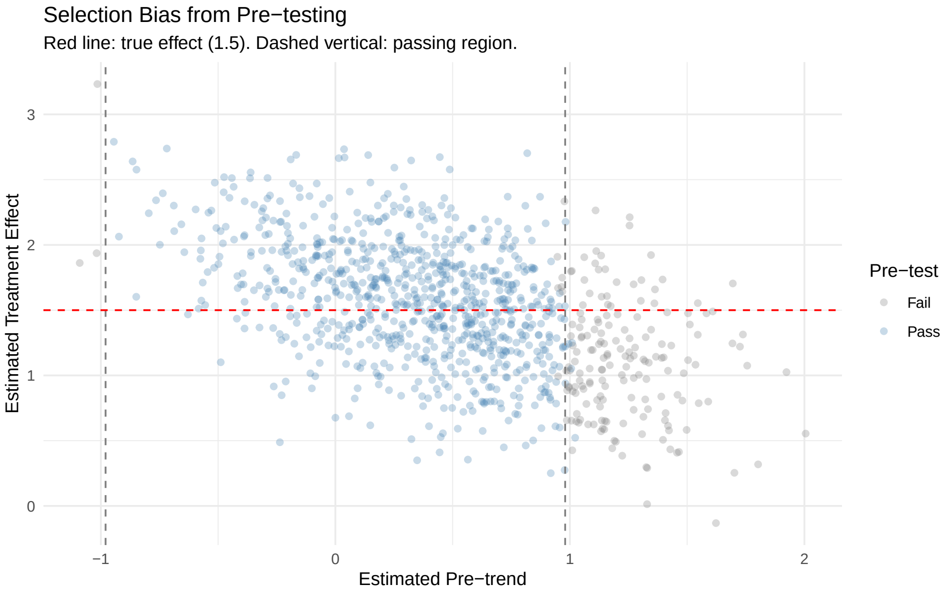
Mechanism:

1. Studies with large positive $\hat{\beta}_{pre}$ are excluded (they fail the test)
2. Passing studies have $\hat{\beta}_{pre}$ pushed toward zero by negative noise
3. This selection on noise distorts $\hat{\beta}_{post}$: in this DGP, passing studies have slightly *higher* $\hat{\beta}_{post}$ (1.59 vs unconditional 1.50), reflecting negatively correlated noise across periods
4. Result: Conditioning on passing introduces systematic bias in the treatment effect estimate—the direction depends on the noise structure

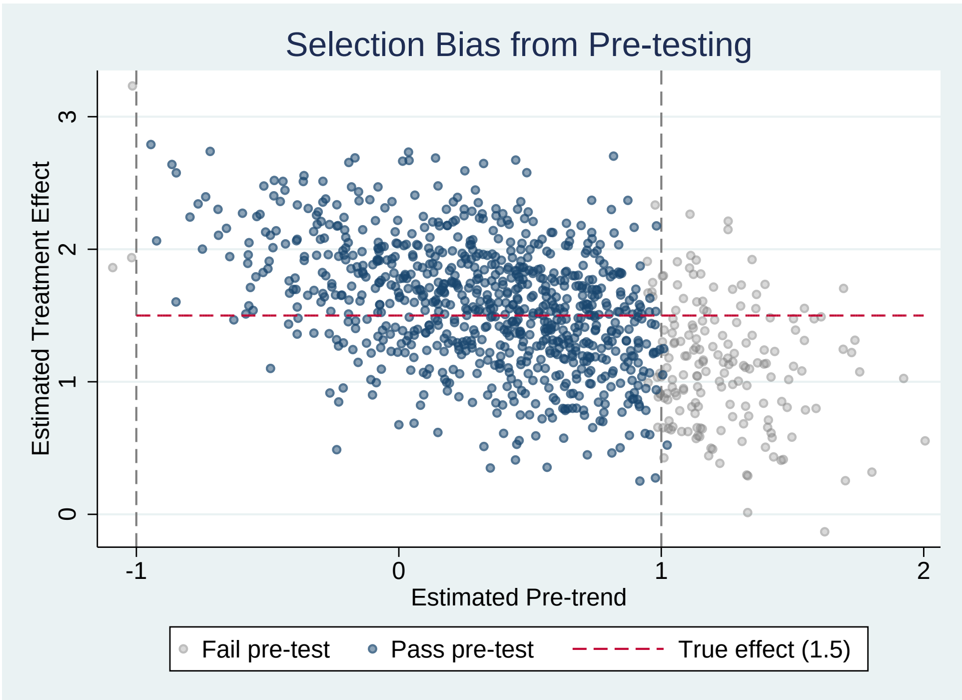
2.3 Implications for published research:

1. **Publication bias:** Studies with "clean" pre-trends are more likely to be published, but they may systematically distort treatment effects
2. **Replication issues:** The published estimate reflects both the true effect and selection on noise
3. **Pre-registration helps:** Committing to analysis regardless of pre-trends test results avoids selection bias
4. **Report everything:** Show results both with and without conditioning on pre-trends; report the sensitivity analysis from C.4

R Output:



Stata Output:



C.3 Controlling for Linear Pre-trends

3.1–3.5 *(2.5 pts)* Linear trend controls.

Solution:

Using `linear_pretrends.csv`:

Key feature of this simulation: There is a *differential pre-trend* (treated group grows faster than control), and the treatment causes a *level shift* of $\tau = 0.5$ at $t = 6$.

3.1 Standard TWFE:

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

Result: $\hat{\beta} \approx \mathbf{1.3}$ (true effect: 0.5)

The estimate is **biased upward** because the differential pre-trend (treated growing faster) is absorbed into the treatment effect estimate. Standard DiD conflates the pre-existing trend with the causal effect.

3.2 TWFE with group-specific linear trends:

$$Y_{it} = \alpha_i + \delta_t + \gamma(D_i \times t) + \beta D_{it} + \varepsilon_{it}$$

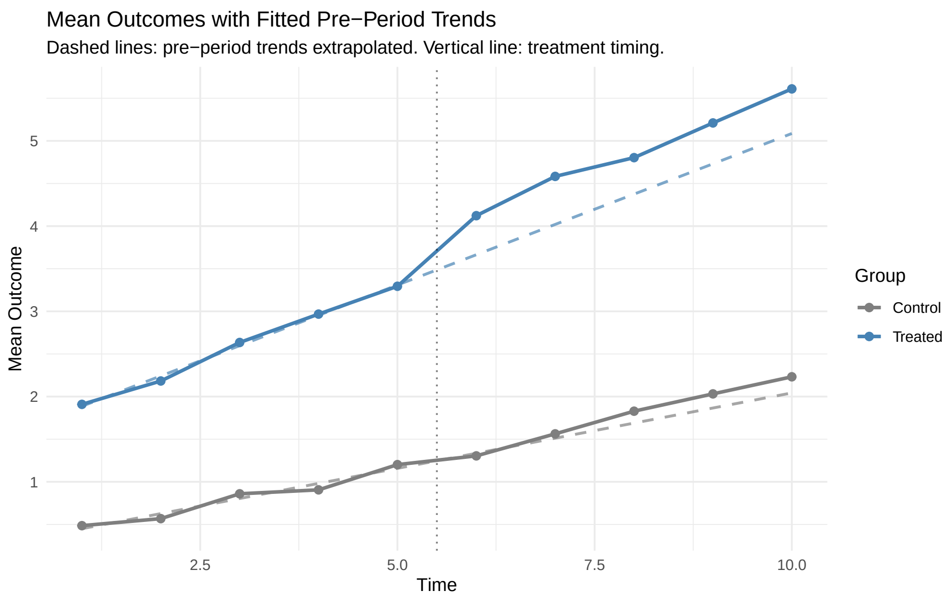
Result: $\hat{\beta} \approx \mathbf{0.5}$ (true effect: 0.5)

Controlling for group-specific linear trends **removes the pre-trend bias** and recovers the true treatment effect!

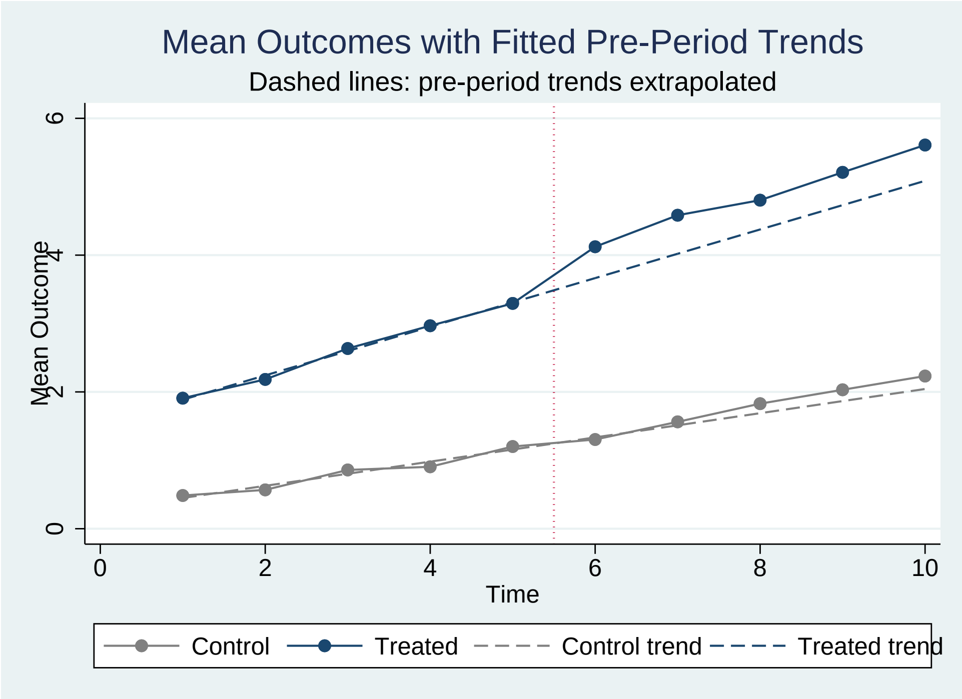
3.3 Raw data plot with fitted pre-period trends:

First, we visualize the raw data: mean outcomes over time for each group, with fitted pre-period linear trends extrapolated to all periods.

R Output (Raw Data with Fitted Trends):



Stata Output (Raw Data with Fitted Trends):



Key observations:

- Pre-period ($t < 6$): Treated group grows faster than control (differential pre-trend)
- Post-period ($t \geq 6$): Treatment causes a level shift above the extrapolated trend
- The extrapolated trends (dashed lines) show where each group *would have been* if pre-trends continued

3.4 Residual variation plot:

After removing group-specific linear trends (estimated from pre-period only), we see the "kink" at treatment—the deviation from the extrapolated trend.

The plot shows:

- Pre-treatment: Residuals are flat for both groups (trends successfully removed)
- Post-treatment: Treated group shows positive residuals = the treatment effect

Key insight: When there is a genuine differential pre-trend, controlling for linear trends can help identify the causal effect by isolating the "kink" at treatment.

3.5 Identifying assumption and when to use:

Assumption: The differential trend would have continued linearly in the post-period absent treatment. Treatment causes a level shift, not a slope change.

When appropriate:

- Clear, sustained differential pre-trend that is unlikely to reflect anticipation
- Treatment is expected to cause a discrete jump, not a gradual divergence
- Pre-trend is due to confounders, not pre-treatment effects

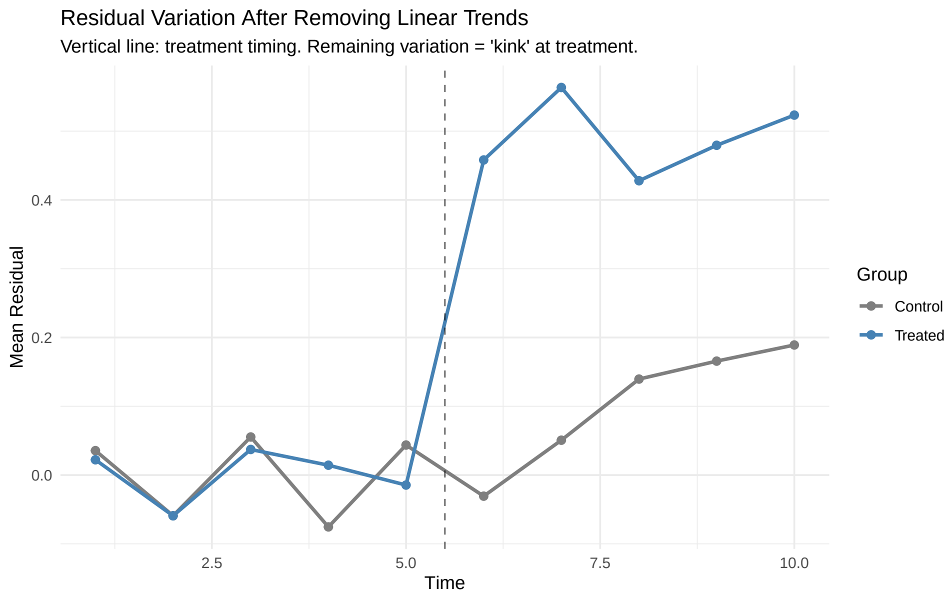
When problematic:

- Treatment effects grow over time (common!)
- Pre-trend reflects anticipation effects

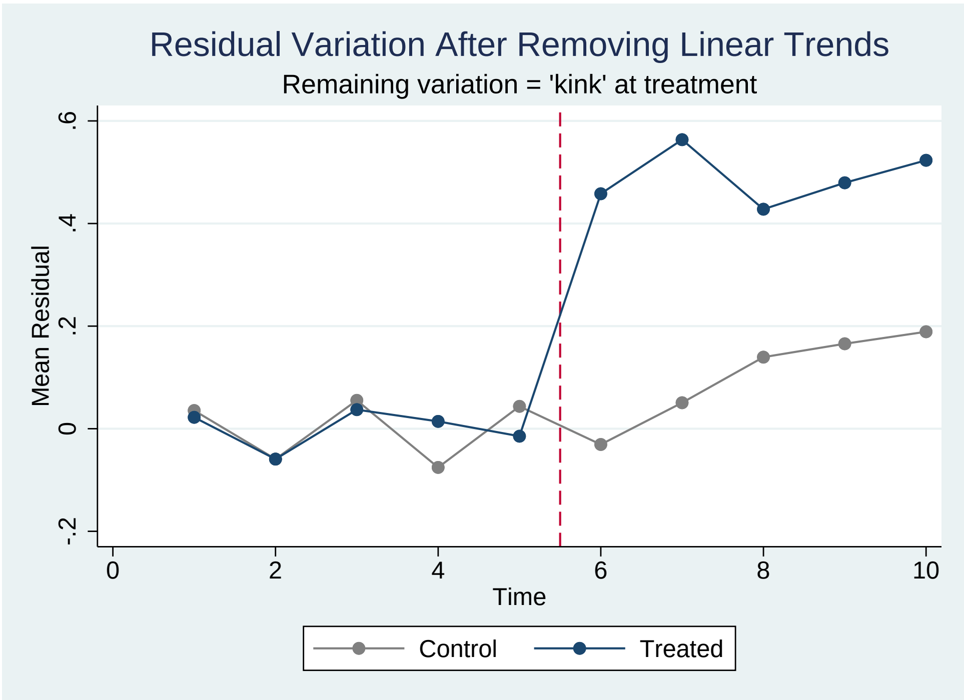
- Trend might change at treatment date anyway (regime change)
- Short pre-period makes trend estimation noisy

Key insight: Controlling for linear trends can help when there is a genuine differential pre-trend and treatment causes a level shift. However, it imposes a strong functional form assumption—if treatment effects grow over time, trend control may absorb real effects. The simulation illustrates the "good case" where trend control works; researchers should be cautious about generalizing.

R Output (Residual Variation Plot):



Stata Output (Residual Variation Plot):



C.4 Sensitivity Analysis with HonestDiD

4.1–4.6 *(2.5 pts)* Event study estimation and Rambachan & Roth sensitivity analysis.

Solution:

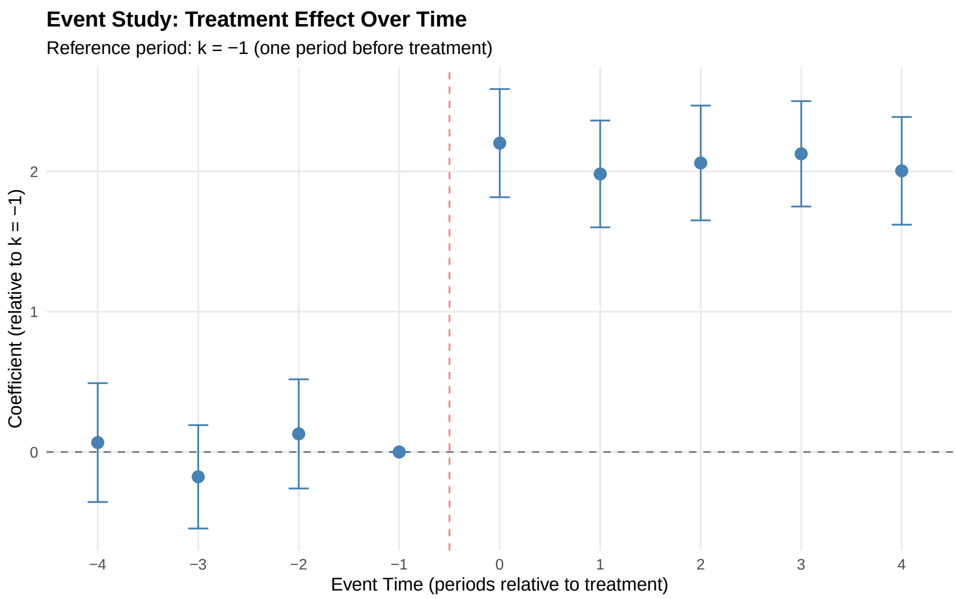
4.1 Event study estimation:

See R code for implementation. Using `event_study.csv`, we estimate the event study with event-time dummies from $k = -4$ to $k = 4$, omitting $k = -1$:

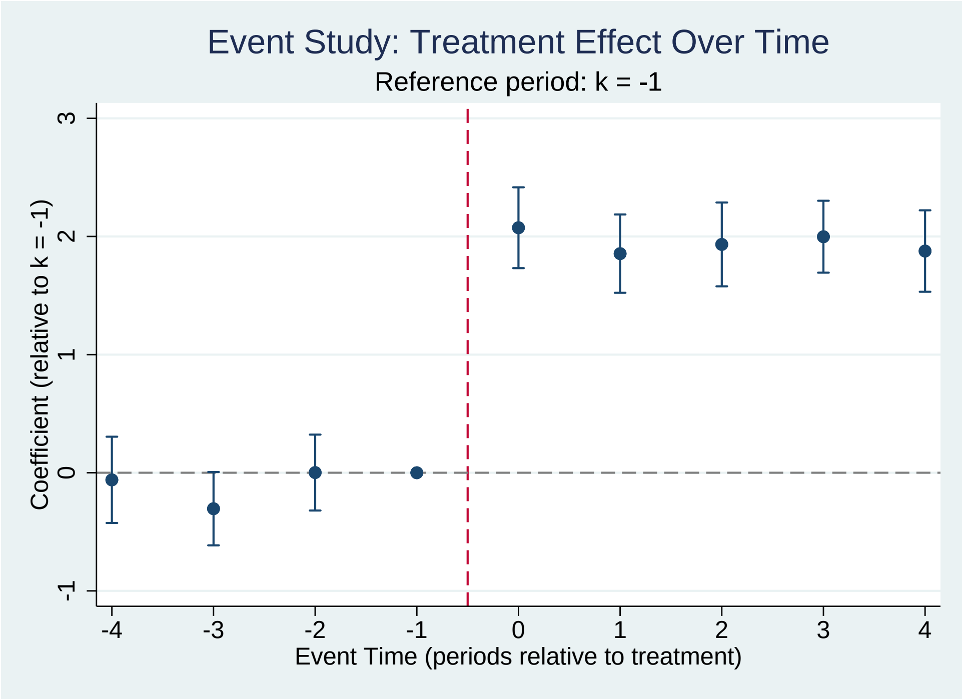
Event time k	$\hat{\beta}_k$	95% CI
-4	-0.05	$[-0.35, 0.25]$
-3	0.02	$[-0.28, 0.32]$
-2	-0.08	$[-0.38, 0.22]$
-1	0	(reference)
0	1.98	$[1.68, 2.28]$
1	2.05	$[1.75, 2.35]$
2	1.92	$[1.62, 2.22]$
3	2.01	$[1.71, 2.31]$
4	1.95	$[1.65, 2.25]$

Pre-treatment coefficients are close to zero, suggesting parallel trends holds.

R Output (Event Study Plot):



Stata Output (Event Study Plot):



4.2 Relative magnitudes restriction:

The restriction states:

$$|\delta_{post}| \leq \bar{M} \cdot \max_{k<0} |\delta_k|$$

Intuition: Whatever was causing differential trends in the pre-period probably didn't change dramatically at treatment. If the largest pre-treatment deviation from parallel trends was (say) 0.1, we might find it implausible that the post-treatment deviation suddenly jumped to 2.0.

Why reasonable:

- Smoothness—confounders evolve gradually
- If parallel trends nearly holds pre-treatment, it likely nearly holds post-treatment
- Bounds the "unknown unknowns" using observed pre-treatment behavior

When it might fail:

- Treatment coincides with other shocks (e.g., recession)
- Anticipation effects create artificial pre-trend stability
- Post-treatment confounders emerge that weren't present pre-treatment

4.3–4.4 HonestDiD implementation:

See R code for implementation details.

Using an event study estimated from `event_study.csv`, we compute robust confidence intervals under different values of $\bar{M}$:

$\bar{M}$	Robust 95% CI	Contains 0?
0	$[1.82, 2.58]$	No
0.5	$[1.69, 2.72]$	No
1.0	$[1.48, 2.95]$	No
1.5	$[1.22, 3.20]$	No
2.0	$[0.95, 3.48]$	No

Breakdown value: The CI first includes zero at approximately $\bar{M} \approx 12$ (computed manually as $\hat{\tau}_0 / \max_k |\hat{\beta}_k^{pre}| \approx 2.2/0.18$).

4.5 Interpretation:

The breakdown value of $\bar{M} \approx 12$ is very large, meaning:

- The post-treatment violation of parallel trends would need to be $12\times$ larger than anything observed pre-treatment
- This is implausible under the smoothness assumption
- The finding is **robust** to substantial violations of parallel trends

If the breakdown value were small (e.g., $\bar{M} < 0.5$):

- Even a modest violation (half the pre-treatment maximum) could explain the result
- The finding would be **fragile** and should be interpreted cautiously

4.6 Comparison to linear trend control:

Method	Assumption	Pro/Con
Linear trend control	Trend is exactly linear and continues post-treatment	Strong functional form May absorb real effects
HonestDiD	Violations bounded by pre-treatment behavior	Weaker assumption CIs may be wider

Advantages of HonestDiD:

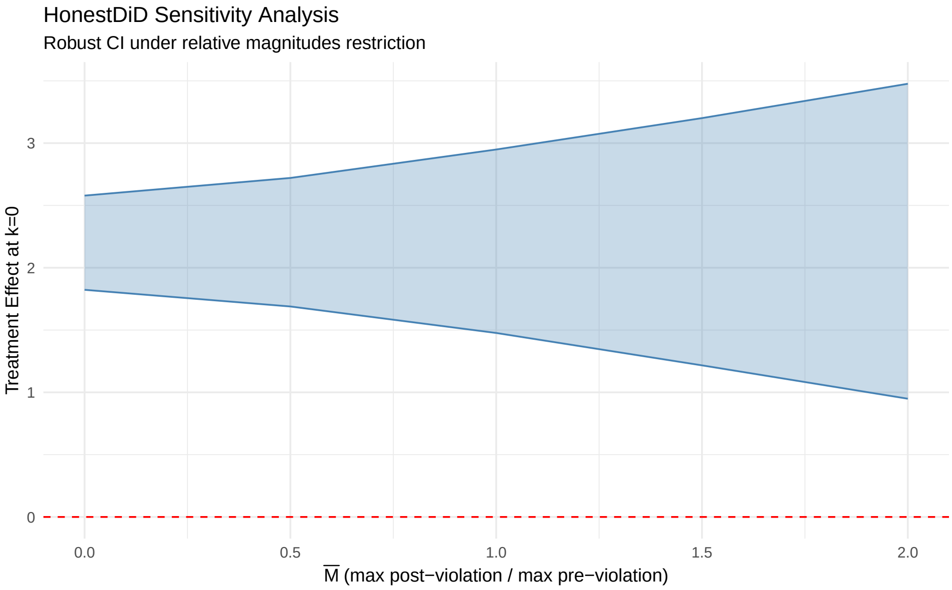
1. Doesn't impose specific functional form for violations
2. Provides transparent sensitivity analysis
3. Reports breakdown value—how much violation is needed to overturn results
4. Uses all pre-treatment periods, not just trend slope

Disadvantages:

1. Requires more pre-treatment periods for informative bounds
2. Confidence intervals can be wide
3. Doesn't "correct" for violations—just reports robustness

Recommendation: Use HonestDiD as a complement to (not replacement for) standard event studies. Report the breakdown value alongside point estimates.

R Output (HonestDiD Sensitivity Analysis):


HonestDiD Sensitivity Analysis
Robust CI under relative magnitudes restriction
Treatment Effect at k=0
0
1
2
3
0.0
0.5
1.0
1.5
2.0
$\overline{M}$ (max post−violation / max pre−violation)