# Lecture 13

## Health // Difference-in-differences

Ivan Rudik
AEM 6510

# Roadmap

- How do we estimate a treatment effect when the treated and control groups do not (counterfactually) look the same in the cross-section?
- What is the mortality cost of lead?

# Difference-in-differences

# Our comparisons so far

So far we've made two kinds of comparisons to estimate treatment effects:

1. Comparing two groups with random assignment to treatment **(RCT)**

# Our comparisons so far

So far we've made two kinds of comparisons to estimate treatment effects:

1. Comparing two groups with random assignment to treatment (RCT)
2. Comparing two groups where there is a local discontinuity (i.e. discrete change) in policy (regression discontinuity)

# Our comparisons so far

So far we've made two kinds of comparisons to estimate treatment effects:

1. Comparing two groups with random assignment to treatment (RCT)
2. Comparing two groups where there is a local discontinuity (i.e. discrete change) in policy (regression discontinuity)

In both of these we are comparing groups in the **cross-section**: there is no concept of time, before and after a policy was enacted, etc

# Our comparisons so far

The key assumption for these comparisons is that

# Our comparisons so far

The key assumption for these comparisons is that the treated group would have looked the same as the control group (i.e. had the same outcomes) in the absence of treatment

# Our comparisons so far

The key assumption for these comparisons is that the treated group would have looked the same as the control group (i.e. had the same outcomes) in the absence of treatment

This assumption is often hard to defend[1]

One way people show that this tends to not be true is to throw in a bunch of extra controls into the regression, if this affects your estimates it indicates there's likely a problem with the assumption.

# Our comparisons so far

The key assumption for these comparisons is that the treated group would have looked the same as the control group (i.e. had the same outcomes) in the absence of treatment

This assumption is often hard to defend[1]

Let's try to relax this assumption by exploiting **temporal comparisons** in addition to the cross-sectional comparison

One way people show that this tends to not be true is to throw in a bunch of extra controls into the regression, if this affects your estimates it indicates there's likely a problem with the assumption.

# Difference-in-differences

One way to describe our comparisons thus far is as **differences**

# Difference-in-differences

One way to describe our comparisons thus far is as **differences**

The estimated effect of a policy is simply the difference in expected outcomes between treated and control groups:

$$\delta = E[Y^1|D = 1] - E[Y^0|D = 0]$$

# Difference-in-differences

One way to describe our comparisons thus far is as **differences**

The estimated effect of a policy is simply the difference in expected outcomes between treated and control groups:

$$\delta = E[Y^1|D = 1] - E[Y^0|D = 0]$$

It's exactly that for an RCT since $D$ was randomly assigned, and its the difference in conditional expectations (conditional on being around the threshold) for RDD

# Difference-in-differences

Our next method is called **difference-in-differences** (DD)

# Difference-in-differences

Our next method is called **difference-in-differences** (DD)

What DD does is take the difference of two comparisons in three steps:

# Difference-in-differences

Our next method is called **difference-in-differences** (DD)

What DD does is take the difference of two comparisons in three steps:

1. Take the difference in mean outcomes between treated and control **before** treatment

# Difference-in-differences

Our next method is called **difference-in-differences** (DD)

What DD does is take the difference of two comparisons in three steps:

1. Take the difference in mean outcomes between treated and control **before** treatment
2. Take the difference in mean outcomes between treated and control **after** treatment

# Difference-in-differences

Our next method is called **difference-in-differences** (DD)

What DD does is take the difference of two comparisons in three steps:

1. Take the difference in mean outcomes between treated and control **before** treatment
2. Take the difference in mean outcomes between treated and control **after** treatment
3. Take the difference between 1 and 2

# Difference-in-differences

Our next method is called **difference-in-differences** (DD)

What DD does is take the difference of two comparisons in three steps:

1. Take the difference in mean outcomes between treated and control **before** treatment
2. Take the difference in mean outcomes between treated and control **after** treatment
3. Take the difference between 1 and 2

The name comes from the fact that we are taking the difference (3) between two differences (1 and 2)

# Difference-in-differences

Why do we use DD?

# Difference-in-differences

Why do we use DD?

The identifying assumption required is less strict than for difference approaches

# Difference-in-differences

Why do we use DD?

The identifying assumption required is less strict than for difference approaches

**DD assumption:** the treatment and control group would have followed **parallel trends** in the absence of treatment

- i.e. the difference in outcomes would have remained constant

# Difference-in-differences

Why do we use DD?

The identifying assumption required is less strict than for difference approaches

**DD assumption:** the treatment and control group would have followed **parallel trends** in the absence of treatment

- i.e. the difference in outcomes would have remained constant

This is much less stringent than requiring the outcomes to have been the same in the absence of treatment

# Difference-in-differences

**Example:** Suppose we want to understand the effect of a conservation policy passed in New York on biodiversity

# Difference-in-differences

**Example:** Suppose we want to understand the effect of a conservation policy passed in New York on biodiversity

Suppose also that:

- The effect of the New York policy is given by B

# Difference-in-differences

**Example:** Suppose we want to understand the effect of a conservation policy passed in New York on biodiversity

Suppose also that:

- The effect of the New York policy is given by B
- Each state has it's own fixed determinants of biodiversity (e.g. land cover, average temperature, etc) given by NY, PA, MA, etc

# Difference-in-differences

**Example:** Suppose we want to understand the effect of a conservation policy passed in New York on biodiversity

Suppose also that:

- The effect of the New York policy is given by B
- Each state has it's own fixed determinants of biodiversity (e.g. land cover, average temperature, etc) given by NY, PA, MA, etc
- Each period has it's own determinants of biodiversity, common across all states (e.g. federal policy, global climate change) given by $T_0$, $T_1$, where 0 is years before the policy is passed, and 1 is after

# Difference-in-differences

When we observe data on biodiversity we see the combination of all determinants: B + NY + T, not just B

# Difference-in-differences

When we observe data on biodiversity we see the combination of all determinants: B + NY + T, not just B

We want to find a way to recover **only** E[B]

# Difference-in-differences

When we observe data on biodiversity we see the combination of all determinants: B + NY + T, not just B

We want to find a way to recover **only** E[B]

There are two ways you could think about trying to estimate B using differences:

# Difference-in-differences

When we observe data on biodiversity we see the combination of all determinants: B + NY + T, not just B

We want to find a way to recover **only** E[B]

There are two ways you could think about trying to estimate B using differences:

1. Compare New York to another state after the policy is passed

# Difference-in-differences

When we observe data on biodiversity we see the combination of all determinants: B + NY + T, not just B

We want to find a way to recover **only** E[B]

There are two ways you could think about trying to estimate B using differences:

1. Compare New York to another state after the policy is passed
2. Compare New York to itself, before and after the policy is passed

# The cross-sectional difference

Let's compare New York to another state, Pennsylvania

# The cross-sectional difference

Let's compare New York to another state, Pennsylvania

If we were to do this with differences we would get an estimate of B given by:

$$(B + NY + T_1) - (PA + T_1) = B + NY - PA$$

# The cross-sectional difference

Let's compare New York to another state, Pennsylvania

If we were to do this with differences we would get an estimate of B given by:

$$(B + NY + T_1) - (PA + T_1) = B + NY - PA$$

This is not B!

# The cross-sectional difference

Let's compare New York to another state, Pennsylvania

If we were to do this with differences we would get an estimate of B given by:

$$(B + NY + T_1) - (PA + T_1) = B + NY - PA$$

This is not B!

Why?

# The cross-sectional difference

$$(B + NY + T_1) - (PA + T_1) = B + NY - PA$$

There are other determinants of biodiversity that are different across New York and Pennsylvania that are **not** the policy: landcover, urbanization, pollution, etc

# The cross-sectional difference

$$(B + NY + T_1) - (PA + T_1) = B + NY - PA$$

There are other determinants of biodiversity that are different across New York and Pennsylvania that are **not** the policy: landcover, urbanization, pollution, etc

If we take a simple difference across states, we can't disentangle whether the difference is due to the policy B or differences in these other factors NY - PA

# The time series difference

The next logical thing to try to circumvent this problem is to compare New York to itself, before $NY + T_0$ and after $B + NY + T_1$ the policy

# The time series difference

The next logical thing to try to circumvent this problem is to compare New York to itself, before $NY + T_0$ and after $B + NY + T_1$ the policy

$$(B + NY + T_1) - (NY + T_0) = B + T_1 - T_0$$

# The time series difference

The next logical thing to try to circumvent this problem is to compare New York to itself, before $NY + T_0$ and after $B + NY + T_1$ the policy

$$(B + NY + T_1) - (NY + T_0) = B + T_1 - T_0$$

This is not $B$!

# The time series difference

The next logical thing to try to circumvent this problem is to compare New York to itself, before $NY + T_0$ and after $B + NY + T_1$ the policy

$$(B + NY + T_1) - (NY + T_0) = B + T_1 - T_0$$

This is not $B$!

Why?

# The time series difference

$$(B + NY + T_1) - (NY + T_0) = B + T_1 - T_0$$

There are other determinants of biodiversity that are different before and after the policy that are **not** the New York policy: federal policy changes, trends in urbanization and pollution

# The time series difference

$$(B + NY + T_1) - (NY + T_0) = B + T_1 - T_0$$

There are other determinants of biodiversity that are different before and after the policy that are **not** the New York policy: federal policy changes, trends in urbanization and pollution

If we take a simple difference over time, we can't disentangle whether the difference is due to the policy $B$ or differences in other factors that are changing over time $T_1 - T_0$

# Difference-in-differences

With DD we **combine** these two differences

# Difference-in-differences

With DD we **combine** these two differences

We take the time series differences for NY and PA:

| | After | Before | Time Series Difference |
|---|---|---|---|
| New York | $B + NY + T_1$ | $NY + T_0$ | $(B + NY + T_1) - (NY + T_0) = B + T_1 - T_0$ |
| Pennsylvania | $PA + T_1$ | $PA + T_0$ | $(PA + T_1) - (PA + T_0) = T_1 - T_0$ |

# Difference-in-differences

With DD we **combine** these two differences

We take the time series differences for NY and PA:

| | After | Before | Time Series Difference |
|---|---|---|---|
| New York | $B + NY + T_1$ | $NY + T_0$ | $(B + NY + T_1) - (NY + T_0) = B + T_1 - T_0$ |
| Pennsylvania | $PA + T_1$ | $PA + T_0$ | $(PA + T_1) - (PA + T_0) = T_1 - T_0$ |

Next, difference the time series differences to get the DD[1]

[1] You can also difference in the opposite order and end up with the same result

# Difference-in-differences

| | After | Before | Time Series Difference |
|---|---|---|---|
| New York | $B + NY + T_1$ | $NY + T_0$ | $(B + NY + T_1) - (B + NY + T_1) = B + T_1 - T_0$ |
| Pennsylvania | $PA + T_1$ | $PA + T_0$ | $(PA + T_1) - (PA + T_0) = T_1 - T_0$ |
| | | Difference-in-differences: | $(B + T_1 - T_0) - (T_1 - T_0) = B$ |

# Difference-in-differences

| | After | Before | Time Series Difference |
|---|---|---|---|
| New York | $B + NY + T_1$ | $NY + T_0$ | $(B + NY + T_1) - (B + NY + T_1) = B + T_1 - T_0$ |
| Pennsylvania | $PA + T_1$ | $PA + T_0$ | $(PA + T_1) - (PA + T_0) = T_1 - T_0$ |
| | | Difference-in-differences: | $(B + T_1 - T_0) - (T_1 - T_0) = B$ |

The time series differences lets us control for all fixed determinants within a state (NY)

# Difference-in-differences

| | After | Before | Time Series Difference |
|---|---|---|---|
| New York | $B + NY + T_1$ | $NY + T_0$ | $(B + NY + T_1) - (B + NY + T_1) = B + T_1 - T_0$ |
| Pennsylvania | $PA + T_1$ | $PA + T_0$ | $(PA + T_1) - (PA + T_0) = T_1 - T_0$ |
| | | Difference-in-differences: | $(B + T_1 - T_0) - (T_1 - T_0) = B$ |

The time series differences lets us control for all fixed determinants within a state (NY)

The cross-sectional difference lets us control for all period-specific determinants common across all states ($T_1$)

# Difference-in-differences

| | After | Before | Time Series Difference |
|---|---|---|---|
| New York | $B + NY + T_1$ | $NY + T_0$ | $(B + NY + T_1) - (B + NY + T_1) = B + T_1 - T_0$ |
| Pennsylvania | $PA + T_1$ | $PA + T_0$ | $(PA + T_1) - (PA + T_0) = T_1 - T_0$ |
| | | Difference-in-differences: | $(B + T_1 - T_0) - (T_1 - T_0) = B$ |

The time series differences lets us control for all fixed determinants within a state (NY)

The cross-sectional difference lets us control for all period-specific determinants common across all states ($T_1$)

Combining these two differences addresses both and lets us recover B, the true effect of the policy!

# Difference-in-differences

Note that DD is not magic

# Difference-in-differences

Note that DD is not magic

It only can address determinants of biodiversity that are either:

# Difference-in-differences

Note that DD is not magic

It only can address determinants of biodiversity that are either:

1. Time-invariant

# Difference-in-differences

Note that DD is not magic

It only can address determinants of biodiversity that are either:

1. Time-invariant
2. Time-varying, but common across all states

# Difference-in-differences

Note that DD is not magic

It only can address determinants of biodiversity that are either:

1. Time-invariant
2. Time-varying, but common across all states

If there is a determinant of biodiversity that is varying over time, and differentially across states, DD will fail to correctly estimate B

- State climate trends, state pollution trends, etc

# Difference-in-differences

Suppose that there is another determinant of biodiversity $C^{NY}{}_1, C^{NY}{}_0$ that only occurs in New York and varies over time

- e.g. climate in New York relative to Pennsylvania

Our DD is then

| | After | Before | Time Series Difference |
|---|---|---|---|
| New York | $B + NY + T_1 + C^{NY}{}_1$ | $NY + T_0 + C^{NY}{}_0$ | $(B + NY + T_1) - (B + NY + T_1) = B + T_1 - T_0 + C^{NY}{}_1 - C^{NY}{}_0$ |
| Pennsylvania | $PA + T_1$ | $PA + T_0$ | $(PA + T_1) - (PA + T_0) = T_1 - T_0$ |
| | | Difference-in-differences: | $(B + T_1 - T_0) - (T_1 - T_0) = B + C^{NY}{}_1 - C^{NY}{}_0$ |

# Difference-in-differences

| | After | Before | Time Series Difference |
|---|---|---|---|
| New York | $B + NY + T_1 + C^{NY}_1 + C^{NY}_0$ | $NY + T_0$ | $(B + NY + T_1) - (B + NY + T_1) = B + T_1 - T_0 + C^{NY}_1 - C^{NY}_0$ |
| Pennsylvania | $PA + T_1$ | $PA + T_0$ | $(PA + T_1) - (PA + T_0) = T_1 - T_0$ |
| | | Difference-in-differences: | $(B + T_1 - T_0) - (T_1 - T_0) = B + C^{NY}_1 - C^{NY}_0$ |

DD cannot isolate the effect of $B$ versus $C^{NY}_1 - C^{NY}_0$

There cannot be any (uncontrolled for) time varying differences between NY and PA if we want to correctly estimate $B$

# Difference-in-differences: to the data

Now lets see how this works in practice: notebook here

# Difference-in-differences: to the data

Now lets see how this works in practice: notebook here

What we are going to do is create a fake dataset where we **know the true value of what we want to estimate**

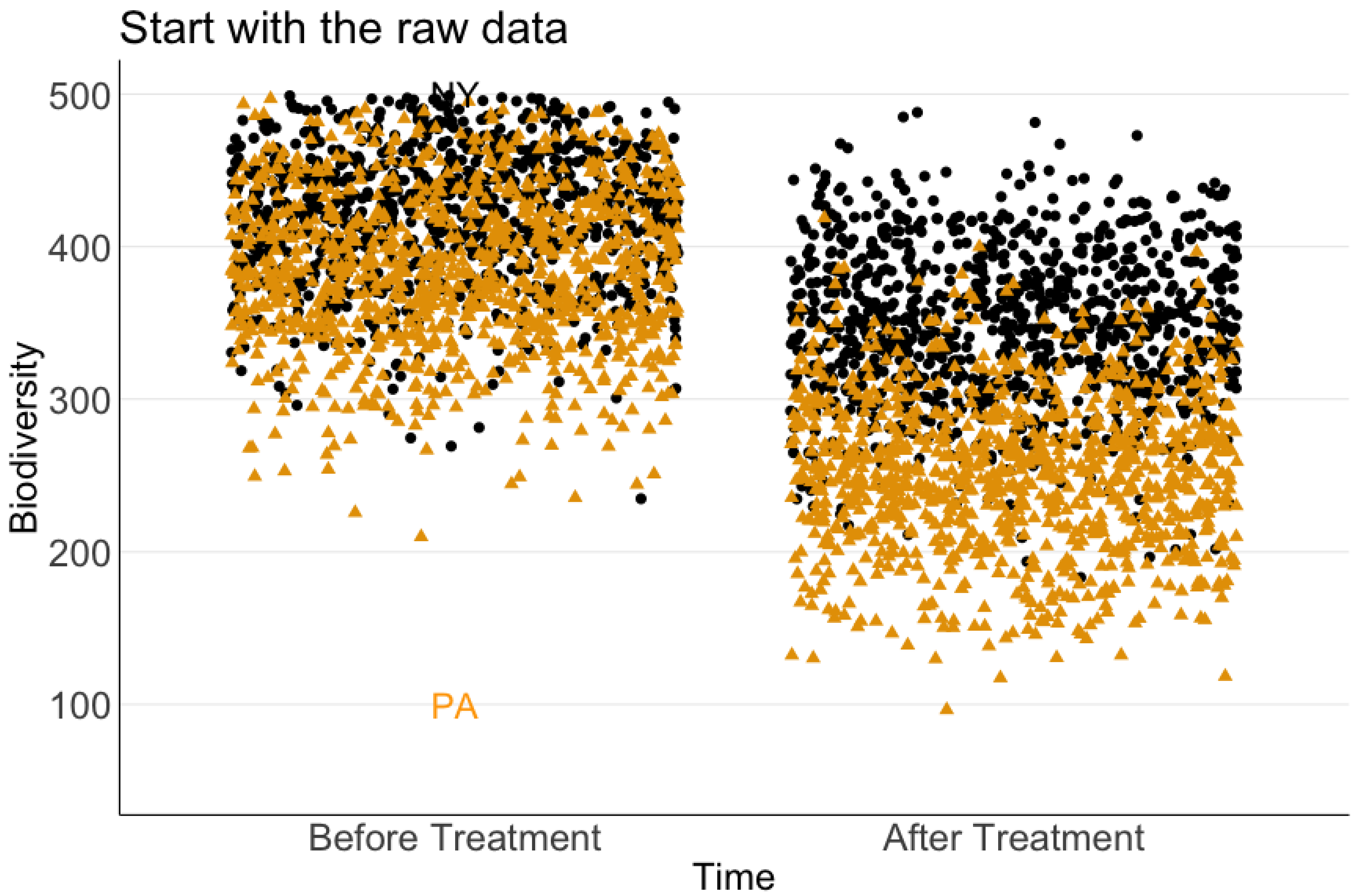
# Difference-in-differences: to the data

Now lets see how this works in practice: notebook here

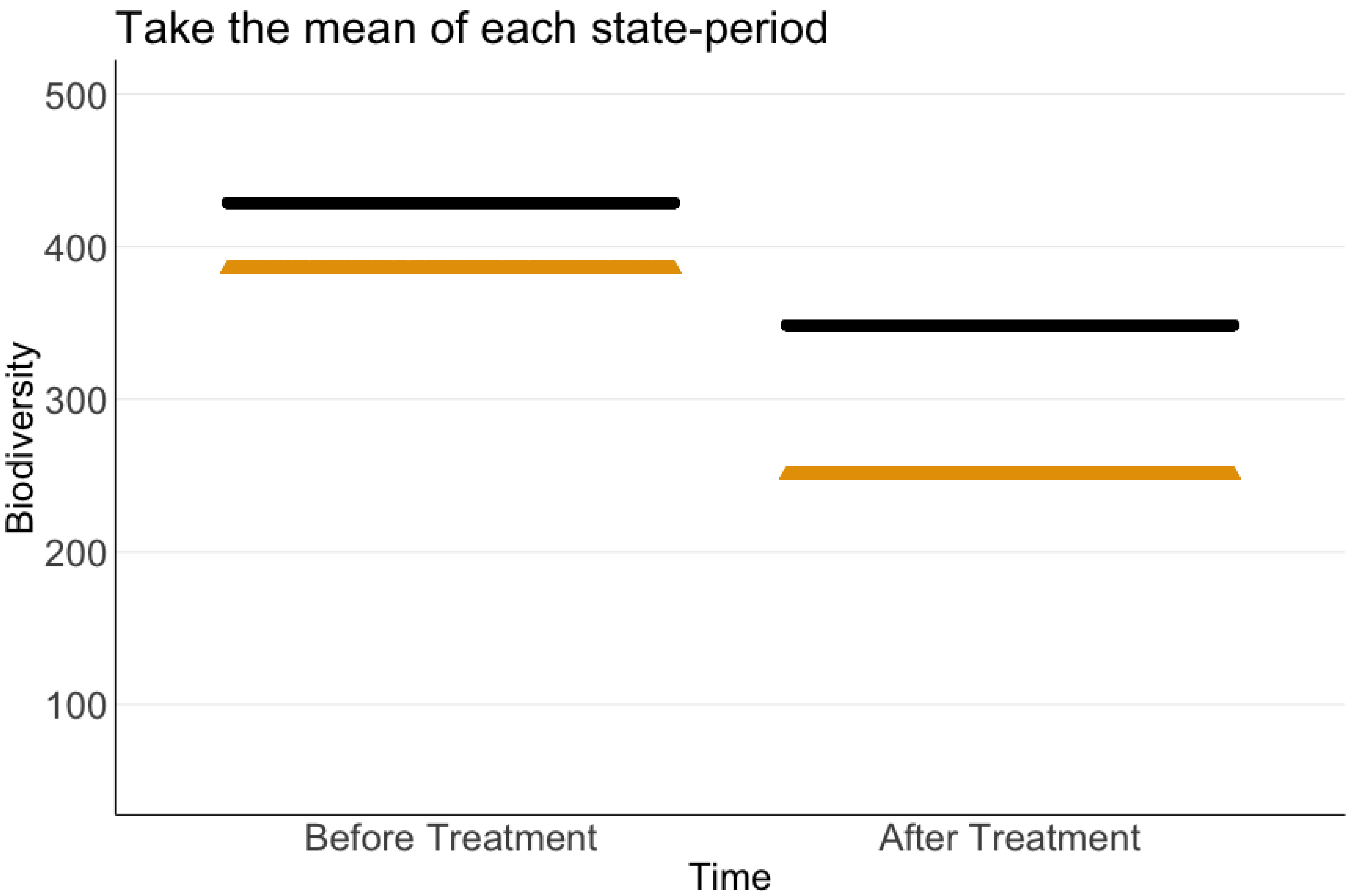
What we are going to do is create a fake dataset where we **know the true value of what we want to estimate**

This is a good skill to practice to make sure you understand methods and that your code works correctly

# DD step 1

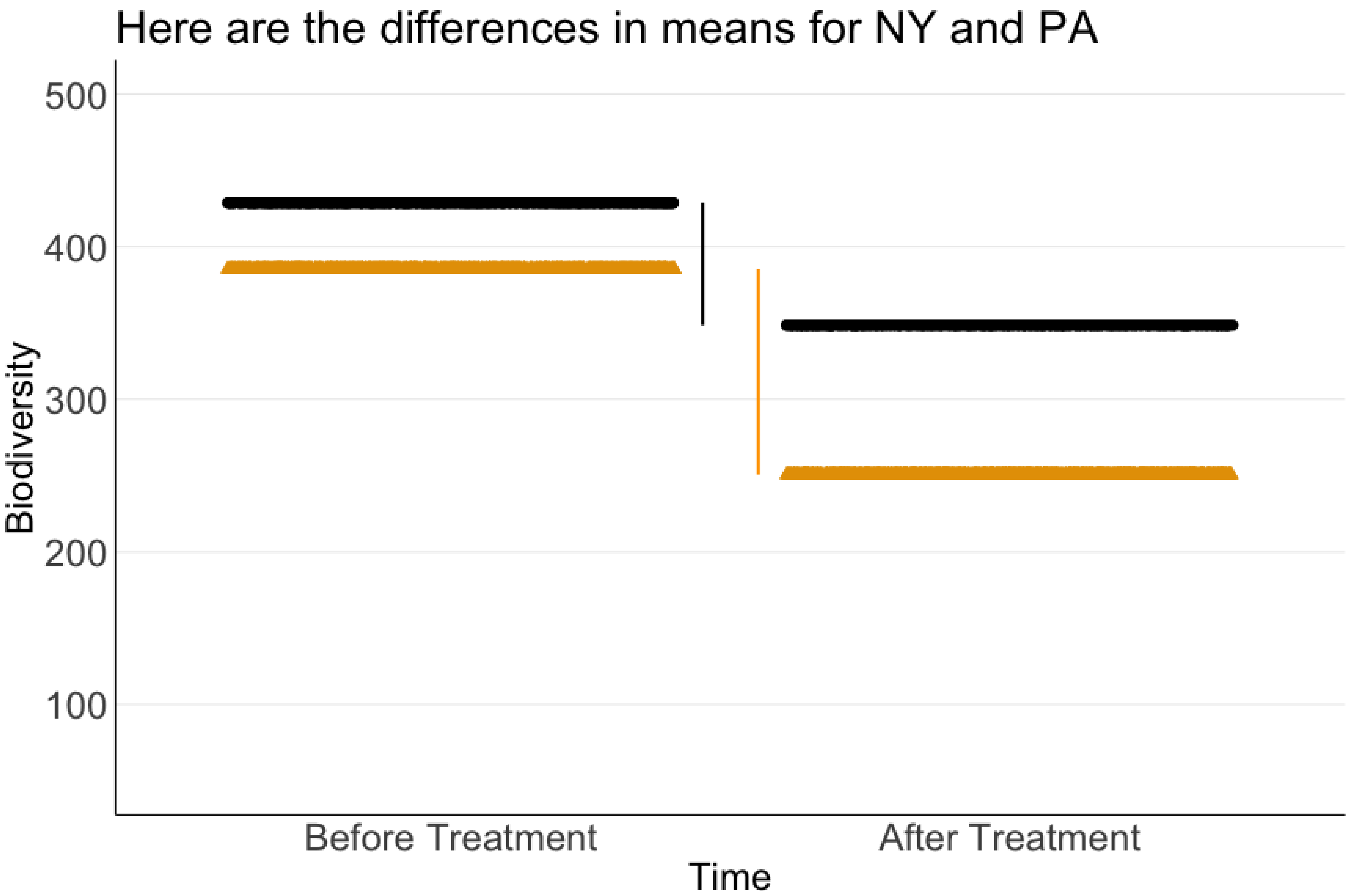

Start with the raw data
NY
PA
500
400
300
200
100
Biodiversity
Before Treatment
After Treatment
Time

# DD step 2



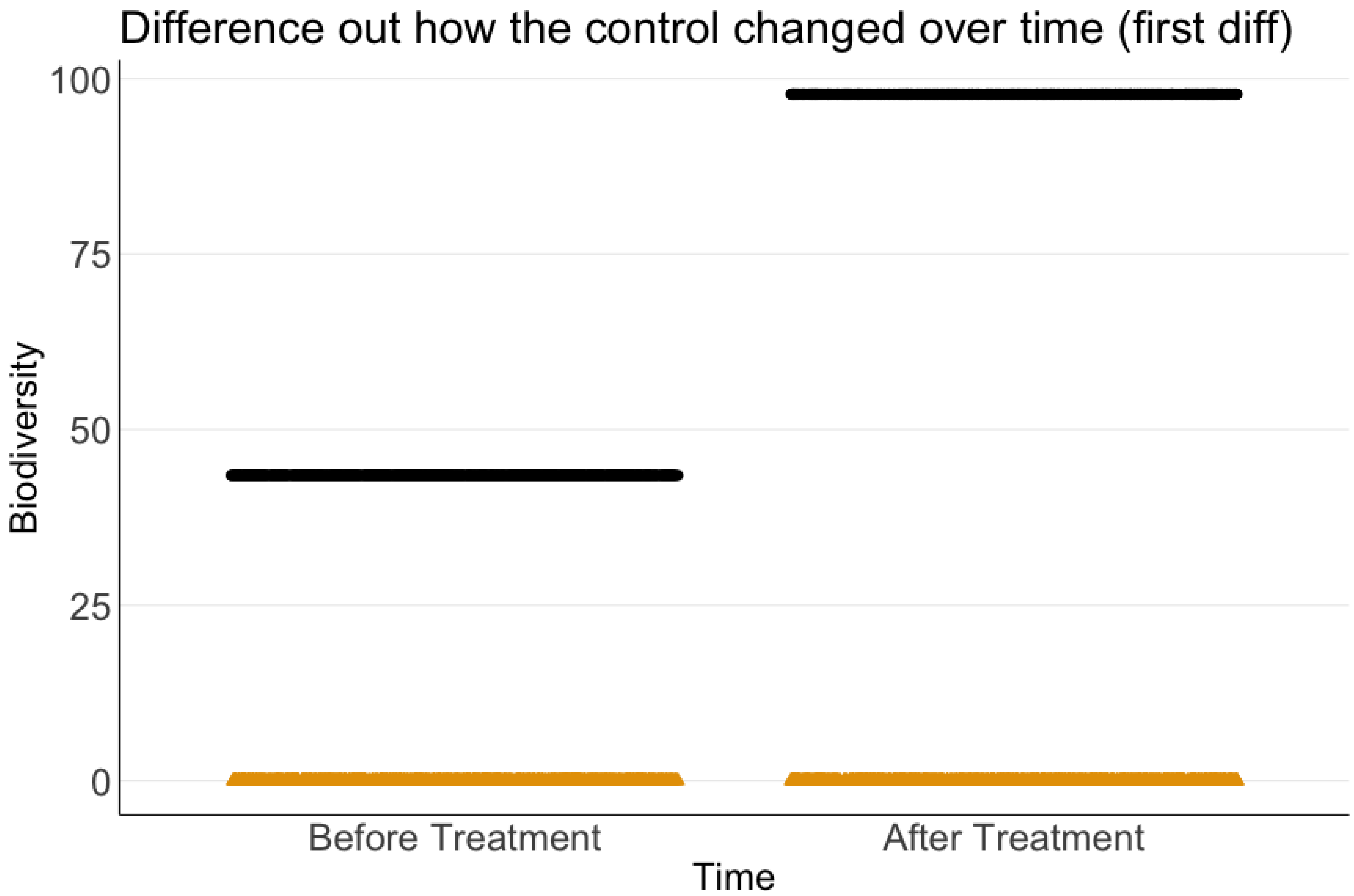
Take the mean of each state-period
500
400
300
200
100
Biodiversity
Before Treatment
After Treatment
Time

# DD step 3



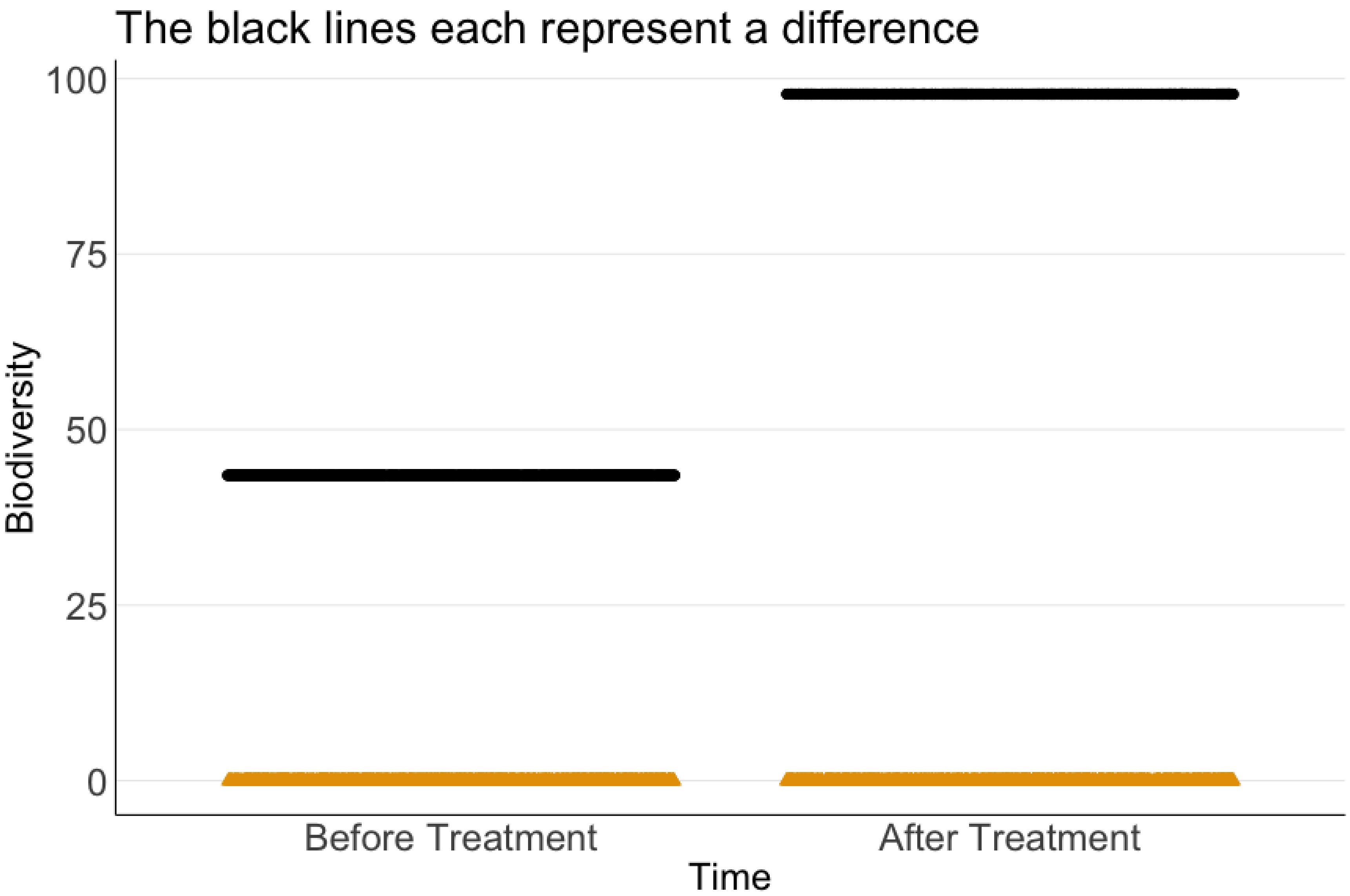
Here are the differences in means for NY and PA
500
400
300
200
100
Biodiversity
Before Treatment
After Treatment
Time

# DD step 4



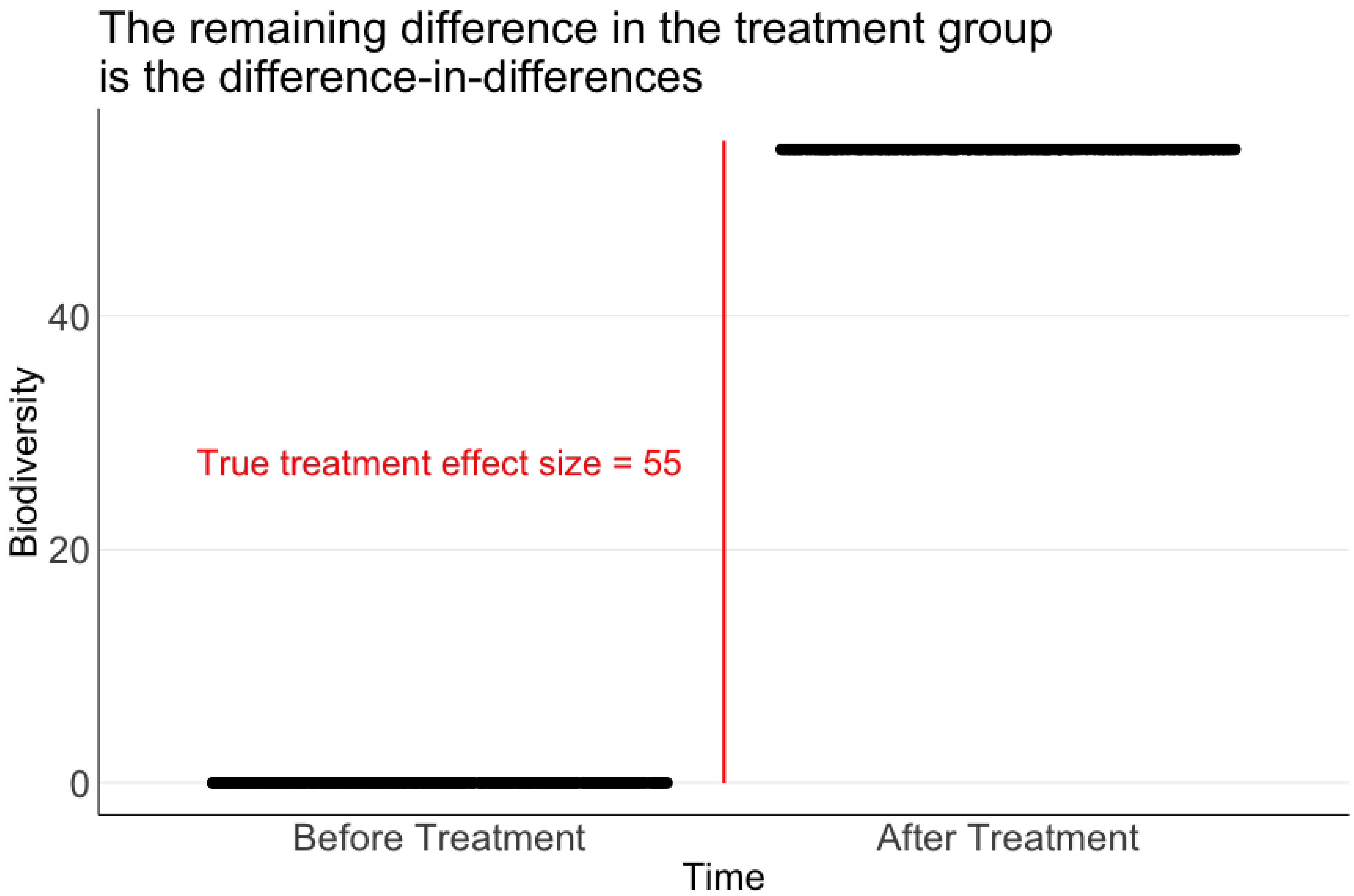
Difference out how the control changed over time (first diff)
100
75
50
25
0
Biodiversity
Before Treatment
After Treatment
Time

# DD step 5



The black lines each represent a difference
100
75
50
25
0
Biodiversity
Before Treatment
After Treatment
Time

# DD step 6


The remaining difference in the treatment group
is the difference-in-differences
True treatment effect size = 55
Biodiversity
40
20
0
Before Treatment
After Treatment
Time

# The effect of leaded gasoline on elderly mortality: Evidence from regulatory exemptions

# What is the paper about?

What is Hollingsworth and Rudik (2021) (HR) about?

# What is the paper about?

What is Hollingsworth and Rudik (2021) (HR) about?

HR aims to estimate the **causal** effect of lead on mortality

# What is the paper about?

What is Hollingsworth and Rudik (2021) (HR) about?

HR aims to estimate the **causal** effect of lead on mortality

Why is this important? We know lead is bad

# What is the paper about?

What is Hollingsworth and Rudik (2021) (HR) about?

HR aims to estimate the **causal** effect of lead on mortality

Why is this important? We know lead is bad

1. There is little causal evidence for any effects of lead

# What is the paper about?

What is Hollingsworth and Rudik (2021) (HR) about?

HR aims to estimate the **causal** effect of lead on mortality

Why is this important? We know lead is bad

1. There is little causal evidence for any effects of lead
2. Almost zero causal evidence for effects of lead on adults in any way

# What is the paper about?

What is Hollingsworth and Rudik (2021) (HR) about?

HR aims to estimate the **causal** effect of lead on mortality

Why is this important? We know lead is bad

1. There is little causal evidence for any effects of lead
2. Almost zero causal evidence for effects of lead on adults in any way
3. Having an accurate measure of effects/costs is vital for policymaking

# Environmental research directly affects policy!

PUBLISHED DOCUMENT

## AGENCY:

Environmental Protection Agency (EPA).

## ACTION:

Notice; call for information.

## SUMMARY:

The Environmental Protection Agency (EPA) is preparing an Integrated Science Assessment (ISA) as part of the review of the primary and secondary National Ambient Air Quality Standards (NAAQS) for Lead (Pb). The ISA will be completed by EPA's Office of Research and Development's Center for Public Health and Environmental Assessment (CPHEA). When final, the ISA is intended to update the previous Pb ISA (EPA/600/R-10/075F), published on June 26, 2013. Interested parties are invited to assist EPA in developing and refining the scientific information base for the review of the Pb NAAQS by submitting research studies and data that have been published, accepted for publication, or presented at a public scientific meeting since January 1, 2011.

DOCUMENT DETAILS

**Printed version:**
PDF

**Publication Date:**
07/07/2020

**Agency:**
Environmental Protection Agency

**Dates:**
All communications and information should be received by EPA by September 8, 2020.

**Document Type:**
Notice

**Document Citation:**
85 FR 40641

**Page:**
40641-40643 (3 pages)

**Agency/Docket Numbers:**
Docket ID No. EPA-HQ-OAR-2020-0312
FRL-10011-92-ORD

**Document Number:**

# How does the paper do it?

HR estimates the causal effect of lead by exploiting a **quasi-experiment**

# How does the paper do it?

HR estimates the causal effect of lead by exploiting a **quasi-experiment**

> a quasi-experiment is a real world occurance that approximates an actual RCT; quasi-experiments are also called natural experiments

# How does the paper do it?

HR estimates the causal effect of lead by exploiting a **quasi-experiment**

> a quasi-experiment is a real world occurance that approximates an actual RCT; quasi-experiments are also called natural experiments

Randomly assigning lead exposure to different groups is unethical, but we can learn from situations where real world exposure was *as good as random*

# How does the paper do it?

HR estimates the causal effect of lead by exploiting a **quasi-experiment**

> a quasi-experiment is a real world occurance that approximates an actual RCT; quasi-experiments are also called natural experiments

Randomly assigning lead exposure to different groups is unethical, but we can learn from situations where real world exposure was *as good as random*

The quasi-experiment HR exploits is the sudden removal of lead from racing gasoline in 2007

# How does the paper do it?

HR estimates the causal effect of lead by exploiting a **quasi-experiment**

> a quasi-experiment is a real world occurance that approximates an actual RCT; quasi-experiments are also called natural experiments

Randomly assigning lead exposure to different groups is unethical, but we can learn from situations where real world exposure was *as good as random*

The quasi-experiment HR exploits is the sudden removal of lead from racing gasoline in 2007

Places that happened to have racetracks in 2007 had a significant decrease in lead emissions relative to places without racetracks

# How does the paper do it?

Here's the 2x2 DD table

| | Before | After |
|---|---|---|
| Treated | Areas near NASCAR racetracks before 2007 | Areas near NASCAR racetracks after 2007 |
| Control | Areas far from NASCAR racetracks before 2007 | Areas far from NASCAR racetracks after 2007 |

We are comparing areas close vs far from racetracks, before vs after deleading in 2007

# How does the paper do it?

Here's the 2x2 DD table

| | Before | After |
|---|---|---|
| Treated | Areas near NASCAR racetracks before 2007 | Areas near NASCAR racetracks after 2007 |
| Control | Areas far from NASCAR racetracks before 2007 | Areas far from NASCAR racetracks after 2007 |

We are comparing areas close vs far from racetracks, before vs after deleading in 2007

# How does the paper do it?

Here's the 2x2 DD table

| | Before | After |
| --- | --- | --- |
| Treated | Areas near NASCAR racetracks before 2007 | Areas near NASCAR racetracks after 2007 |
| Control | Areas far from NASCAR racetracks before 2007 | Areas far from NASCAR racetracks after 2007 |

We are comparing areas close vs far from racetracks, before vs after deleading in 2007

# How does the paper do it?

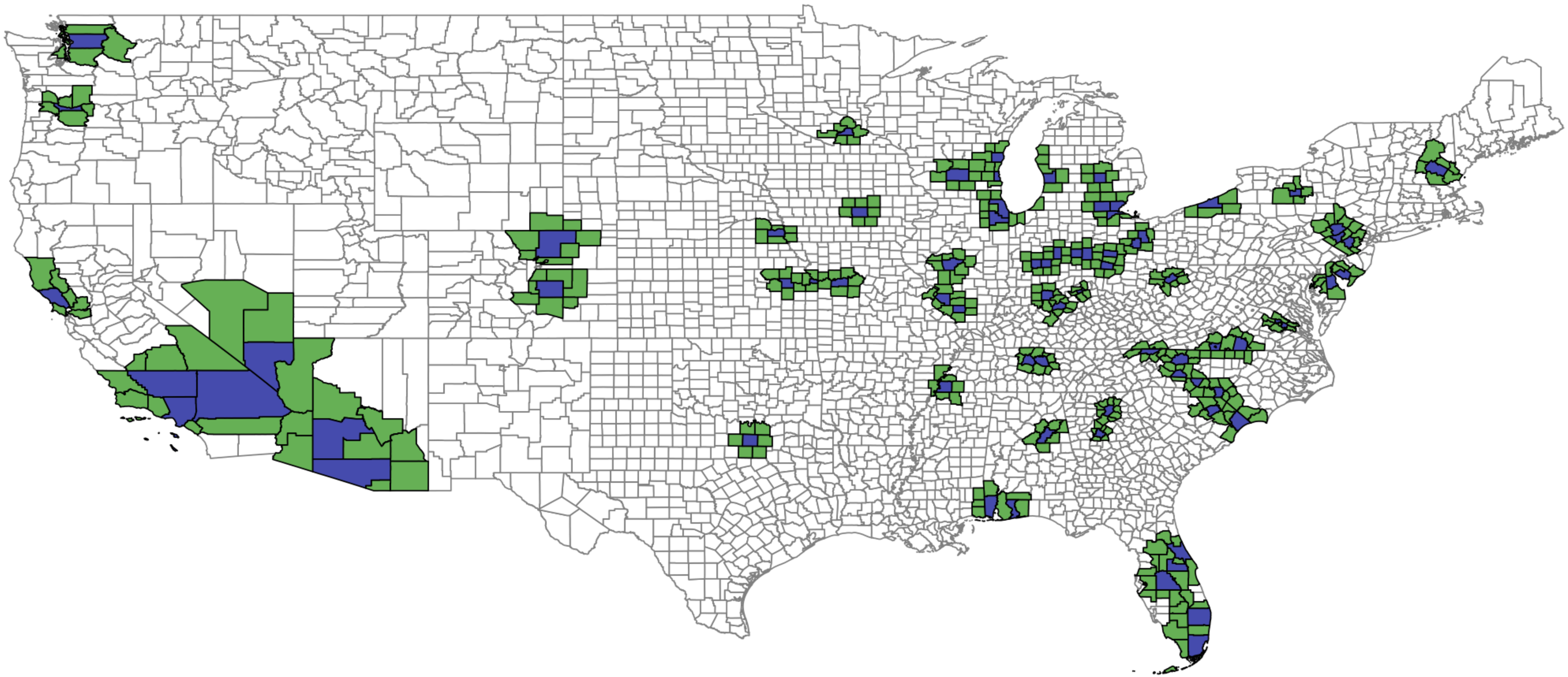
Here's the 2x2 DD table

| | Before | After |
| --- | --- | --- |
| Treated | Areas near NASCAR racetracks before 2007 | Areas near NASCAR racetracks after 2007 |
| Control | Areas far from NASCAR racetracks before 2007 | Areas far from NASCAR racetracks after 2007 |

We are comparing areas close vs far from racetracks, before vs after deleading in 2007

# How does the paper do it?

track/treated counties, border/weak treated counties, white control counties



# Event studies

What HR do is actually a generalization of DD: an **event study**

# Event studies

What HR do is actually a generalization of DD: an **event study**

An event study is basically just a DD with multiple time periods before and after the treatment

# Event studies

What HR do is actually a generalization of DD: an **event study**

An event study is basically just a DD with multiple time periods before and after the treatment

Here we estimate the effect of being in the treated group, relative to some baseline period

- In the 2x2 DD this baseline period was "before"

# Event studies

What HR do is actually a generalization of DD: an **event study**

An event study is basically just a DD with multiple time periods before and after the treatment

Here we estimate the effect of being in the treated group, relative to some baseline period

- In the 2x2 DD this baseline period was "before"

1. We can look at dynamic effects

# Event studies

What HR do is actually a generalization of DD: an **event study**

An event study is basically just a DD with multiple time periods before and after the treatment

Here we estimate the effect of being in the treated group, relative to some baseline period

- In the 2x2 DD this baseline period was "before"

1. We can look at dynamic effects
2. It provides supporting evidence for the parallel trends assumption

# Event studies: dynamic effects

With an event study we can see how the effect of a policy changes *over time*

# Event studies: dynamic effects

With an event study we can see how the effect of a policy changes *over time*

We will be able to estimate the effect in each post-period relative to some period

# Event studies: dynamic effects

With an event study we can see how the effect of a policy changes *over time*

We will be able to estimate the effect in each post-period relative to some period

If we see estimates getting larger, the effect grows over time

# Event studies: dynamic effects

With an event study we can see how the effect of a policy changes *over time*

We will be able to estimate the effect in each post-period relative to some period

If we see estimates getting larger, the effect grows over time

If estimates get smaller, the effect shrinks

# Event studies: parallel pre-trends

In an event study we may have multiple periods before treatment

# Event studies: parallel pre-trends

In an event study we may have multiple periods before treatment

We can estimate the effect of being in the treatment group in each of these (pre-)periods

# Event studies: parallel pre-trends

In an event study we may have multiple periods before treatment

We can estimate the effect of being in the treatment group in each of these (pre-)periods

If there is no trend in the effect, the **pre-**trends are parallel

# Event studies: parallel pre-trends

In an event study we may have multiple periods before treatment

We can estimate the effect of being in the treatment group in each of these (pre-)periods

If there is no trend in the effect, the **pre-**trends are parallel

This gives us some comfort that the trends were likely to be parallel in the post-period in the absence of treatment

# Event studies: parallel pre-trends

In an event study we may have multiple periods before treatment

We can estimate the effect of being in the treatment group in each of these (pre-)periods

If there is no trend in the effect, the **pre-**trends are parallel

This gives us some comfort that the trends were likely to be parallel in the post-period in the absence of treatment

We never observe what would have **actually** happened though, parallel pre-trends is just supporting evidence that the parallel trends assumption holds

# HR (2021): to the data

Now lets use the actual HR (2021) data to get some practice: notebook here

# HR (2021): takeaways

What can we take away from HR (2021)

1. Lead has a causal effect on elderly mortality
   - It appears to be through cardiovascular channels, perhaps *deaths of despair* as well (Case and Deaton, 2015)
2. The mortality costs are large
   - The external mortality costs imposed by NASCAR are larger than the value of all NASCAR teams combined
   - The social cost of lead is **at least $1,000 per gram**
   - This is very, very, very big