# Diff-in-diff I & II

## Sessions 8–9

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

# Plan for today

Quasi-experiments

Interactions & regression

Two wrongs make a right

Diff-in-diff assumptions

# Quasi-experiments

**RCTs are great!**

**Super impractical to do all the time though!**

# Quasi-experiments

You can't always randomly assign people to do things

So let other people (or the government, or nature, or something else) do it for you

# Quasi-experiments

**Quasi-experiment**
**A situation where you, as researcher, did not assign people to treatment/control**

**External validity** 👍 **Selection** 👎

**Assignment to treatment is "as if" random**

# Quasi-experiments vs. DAG adjustment

We did a lot of work with DAGs!
You're good at closing backdoors with matching and IPW

DAGs can work for any kind of observational data,
even without a quasi-experimentalish situation

**Quasi-experiments are a little different:**
**the context isolates pathway between treatment and outcome**

They're wildly popular in social sciences (especially economics!),
maybe more credible (?) there than just making DAG adjustments

You can still draw a DAG for a quasi-experiment though!

# Analyzing quasi-experiments

## Difference-in-differences

**DiD; DD; diff-in-diff**

## Regression discontinuity

**RD; RDD**

## Instrumental variables

**IV**

# Interactions & regression

# Sliders and switches

$$\widehat{\text{Happiness}} = \beta_0 + \beta_1 \text{Life expectancy} + \beta_2 \text{Latin America} + \varepsilon$$

.small-code[

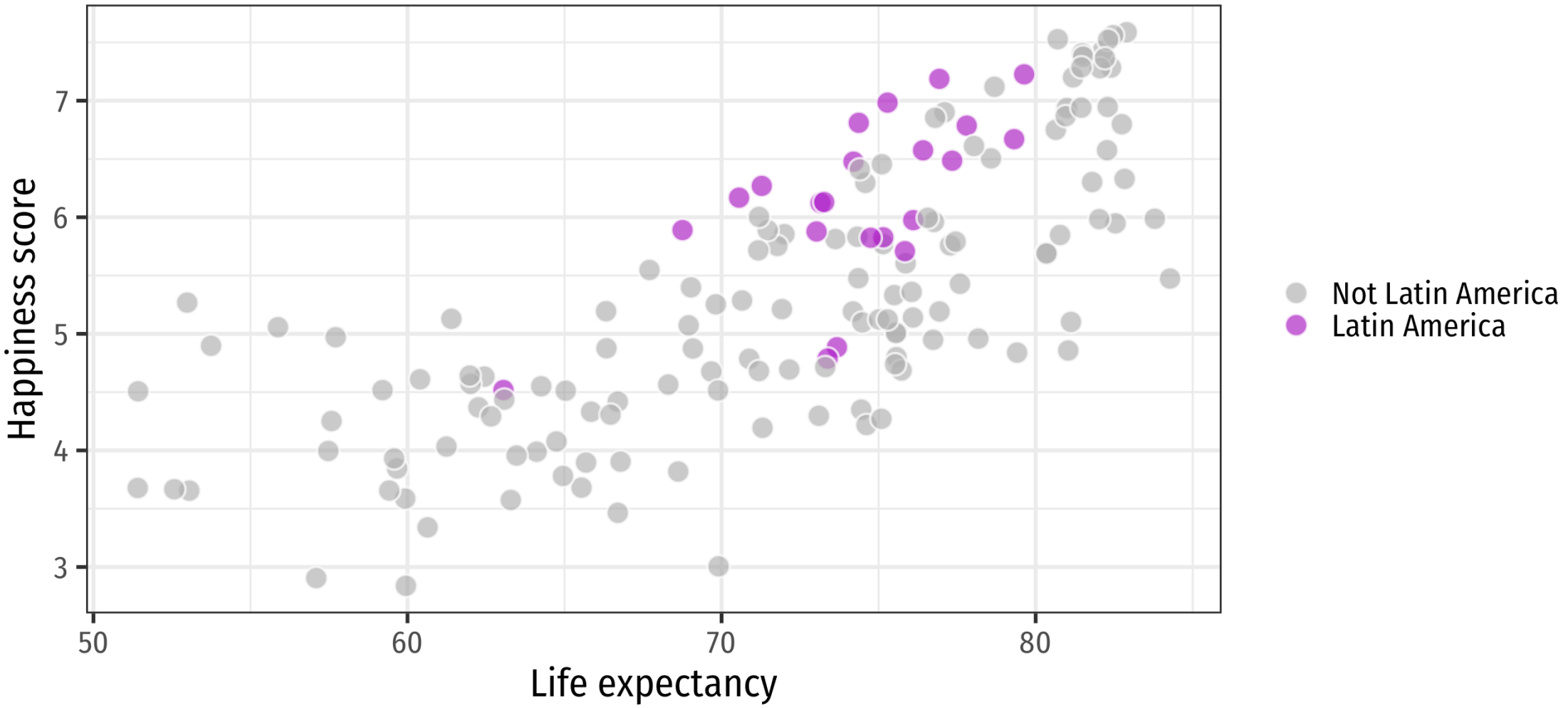
```
model1 <- lm(happiness_score ~ life_expectancy + latin_america
             data = world_happiness)
tidy(model1)
```

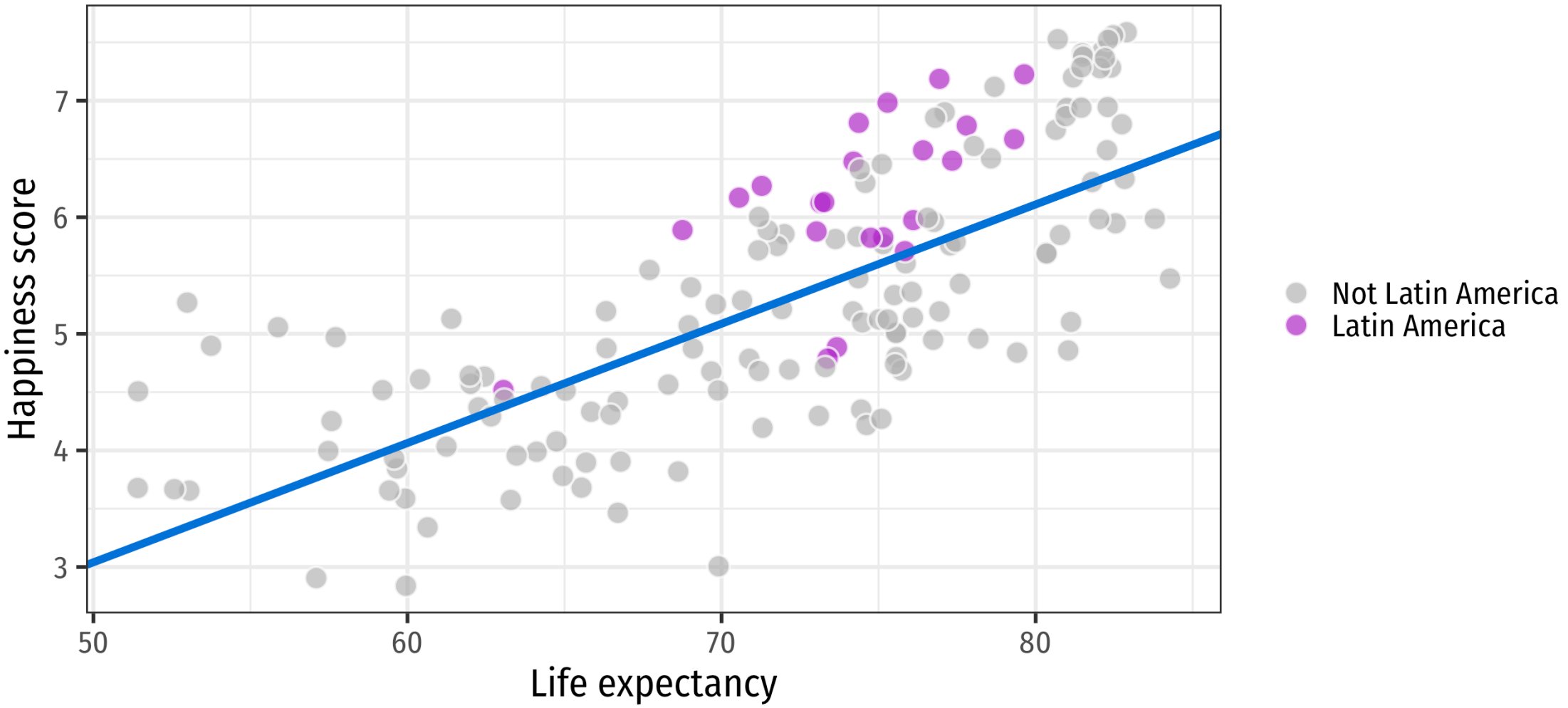
```
##  [38;5;246m# A tibble: 3 × 5 [39m
##    term                          estimate std.error statistic  p.v
##     [3m [38;5;246m<chr> [39m [23m                                [3m [38
##  [38;5;250m1 [39m (Intercept)                 - [31m2 [39m [31m
##  [38;5;250m2 [39m life_expectancy               0.102    0.007 [4
##  [38;5;250m3 [39m latin_americaLatin America    0.623    0.173
```

# Indicators and interactions

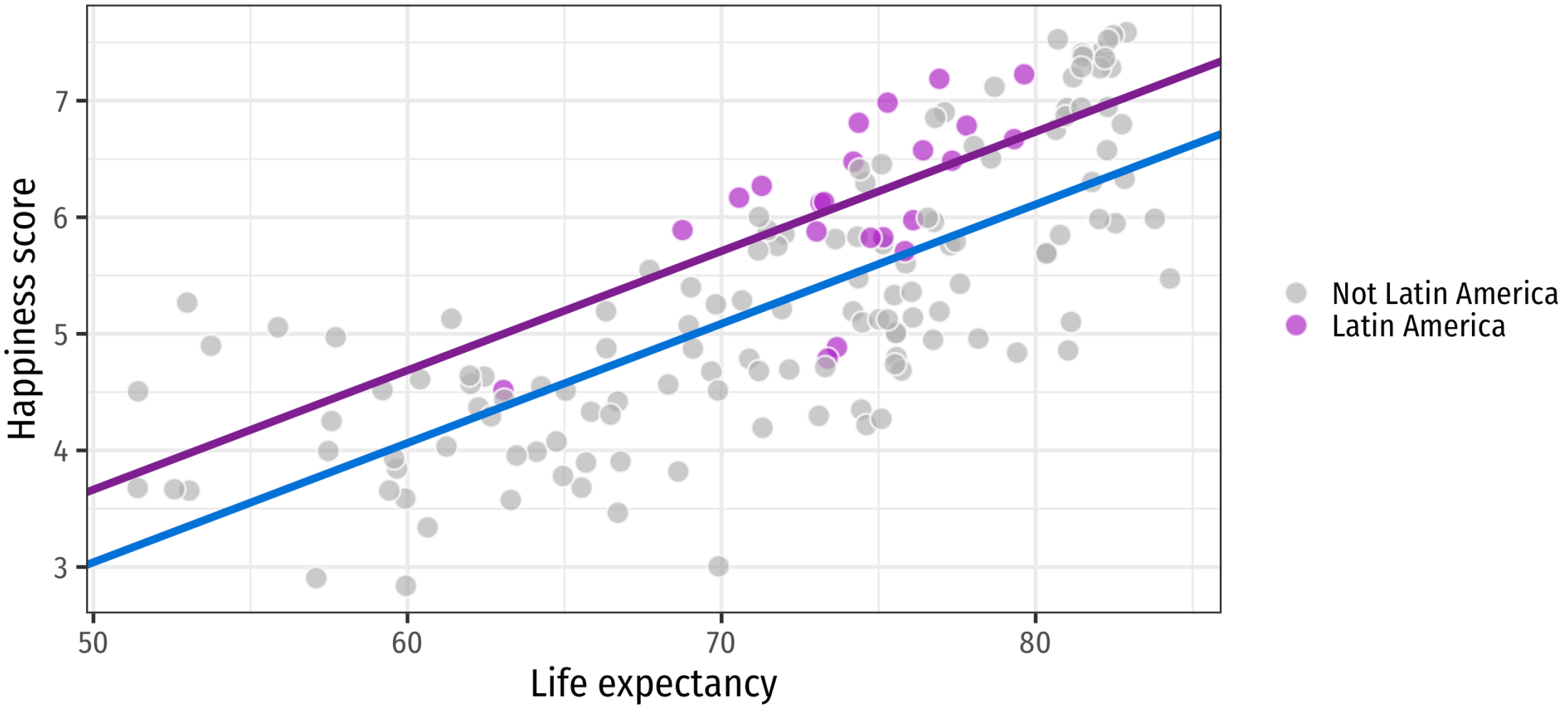
**Indicators** (dummies)

Change in intercept for specific group


Happiness score
7
6
5
4
3
50
60
70
80
Life expectancy
Not Latin America
Latin America




**World slope = 0.102**




**Latin America intercept shifted up 0.62; line has same slope as world (0.102)**

$$\widehat{\text{Happiness}} = \beta_0 + \beta_1 \text{Life expectancy} + \beta_2 \text{Latin America} + \beta_3 (\text{Life expectancy} \times \text{Latin America}) + \varepsilon$$

## .small-code[

```
model2 <- lm(happiness_score ~ life_expectancy + latin_america
             (life_expectancy * latin_america), data = world
tidy(model2)
```

```
##  [38;5;246m# A tibble: 4 × 5 [39m
##   term                                  estimate std.error
##    [3m [38;5;246m<chr> [39m [23m
##  [38;5;250m1 [39m (Intercept)                                - [
##  [38;5;250m2 [39m life_expectancy                             0.
```

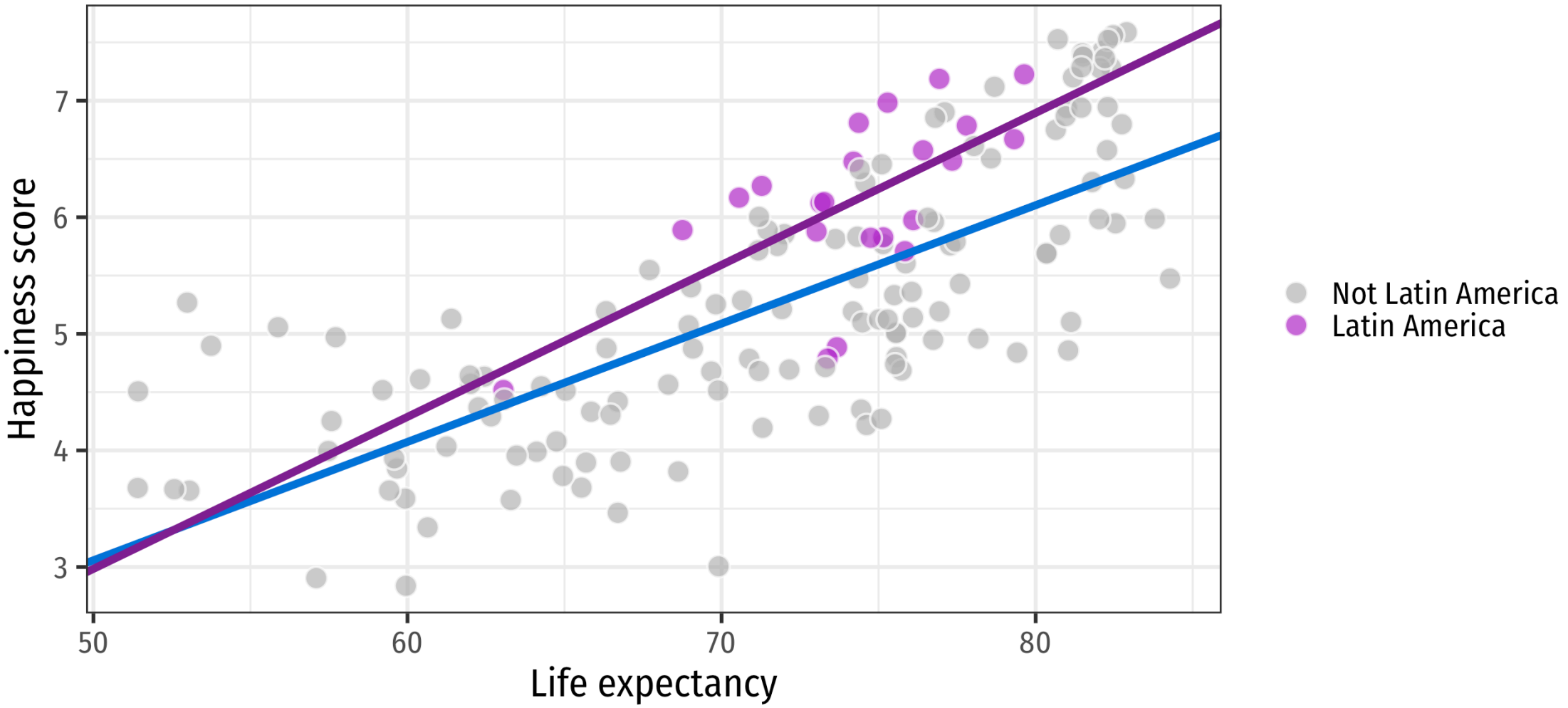
# Indicators and interactions

**Indicators** (dummies)

Change in **intercept** for specific group

**Interactions**

Change in **slope** for specific group




**Latin America slope is 0.029 + 0.102 = 0.13; different from rest of the world**

# Interactions

**What would happen if you ran this?**

```
model3 <- lm(happiness_score ~ (life_expectancy * latin_america),
             data = world_happiness)
```

## .small-code[

```
##  [38;5;246m# A tibble: 4 × 5 [39m
##   term                                    estimate std.error
##    [3m [38;5;246m<chr> [39m [23m
##  [38;5;250m1 [39m (Intercept)                                  - [3
##  [38;5;250m2 [39m life_expectancy                               0.
##  [38;5;250m3 [39m latin_americaLatin America                    - [1
```

# Interactions

**What would happen if you ran this?**

```
model4 <- lm(happiness_score ~ life_expectancy * region, # region has multiple categories
             data = world_happiness)
```

.pull-left-wide[ .small-code.smaller[

```
##  [38;5;246m# A tibble: 14 × 5 [39m
##    term                                    estimate std.error
##    [3m [38;5;246m<chr> [39m [23m                        
##  [38;5;250m 1 [39m (Intercept)                          -
##  [38;5;250m 2 [39m life_expectancy                       0
##  [38;5;250m 3 [39m regionEurope & Central Asia          -
```

# General idea of interactions

**The *additional* change that happens when combining two explanatory variables**

**Life expectancy effect**

**Latin America effect**

**Additional life expectancy effect in Latin America**

**Is there a discount when combining cheese and chili?**

**What is the cheese effect?**

**What is the chili effect?**

**What is the chili × cheese effect?**

# Two wrongs make a right

I
federalism
(for the natural experiments)

# Raising the minimum wage

## What happens if you raise the minimum wage?

**Economic theory says there should be fewer jobs**

## New Jersey in 1992

## $4.25 → $5.05

# Before vs. after

## Average # of jobs per fast food restaurant in NJ

$\text{New Jersey}_{\text{Before change}} = 20.44$

$\text{New Jersey}_{\text{After change}} = 21.03$

$\Delta = 0.59$

## Is this the causal effect?

# Treatment vs. control

## Average # of jobs per fast food restaurant

$\text{Pennsylvania}_{\text{After change}} = 21.17$

$\text{New Jersey}_{\text{After change}} = 21.03$

$\Delta = -0.14$

## Is this the causal effect?

# Problems

## Comparing only before/after

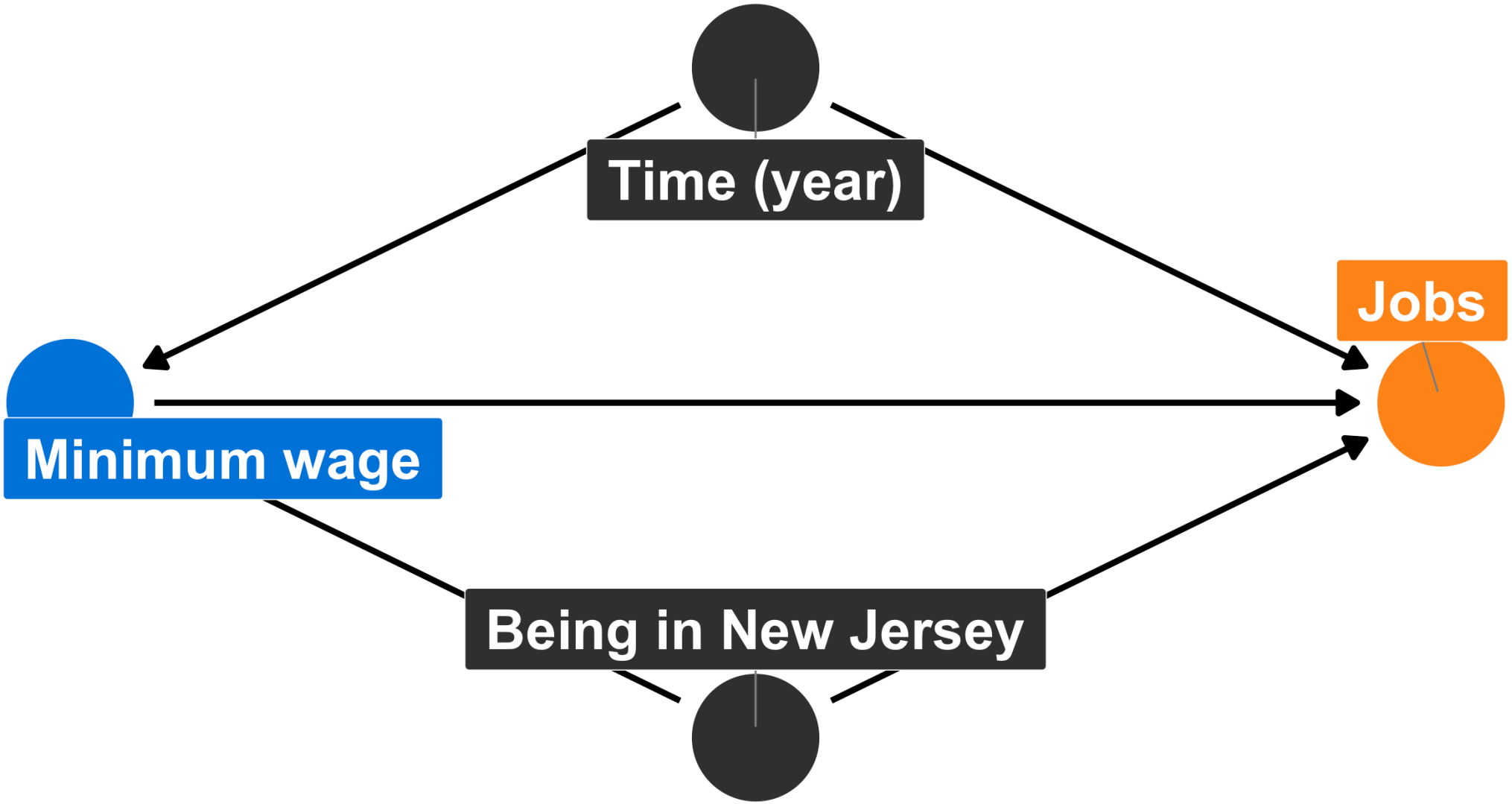
**You're only looking at the treatment group!**

**Impossible to know if change happened because of treatment or just naturally**

## Comparing only treatment/control

**You're only looking at post-treatment values**

**Impossible to know if change happened because of natural growth**


Time (year)
Jobs
Minimum wage
Being in New Jersey

| | Pre mean | Post mean | |
|---|---|---|---|
| **Control** | **A** (never treated) | **B** (never treated) | |
| **Treatment** | **C** (not yet treated) | **D** (treated) | |

| | Pre mean | Post mean | Δ (post − pre) |
|---|---|---|---|
| **Control** | **A** (never treated) | **B** (never treated) | **B − A** |
| **Treatment** | **C** (not yet treated) | **D** (treated) | **D − C** |

**Δ (post − pre) = within-unit growth**

| | Pre mean | Post mean | |
|---|---|---|---|
| **Control** | **A** (never treated) | **B** (never treated) | |
| **Treatment** | **C** (not yet treated) | **D** (treated) | |
| **Δ (treatment – control)** | **C – A** | **D – B** | |

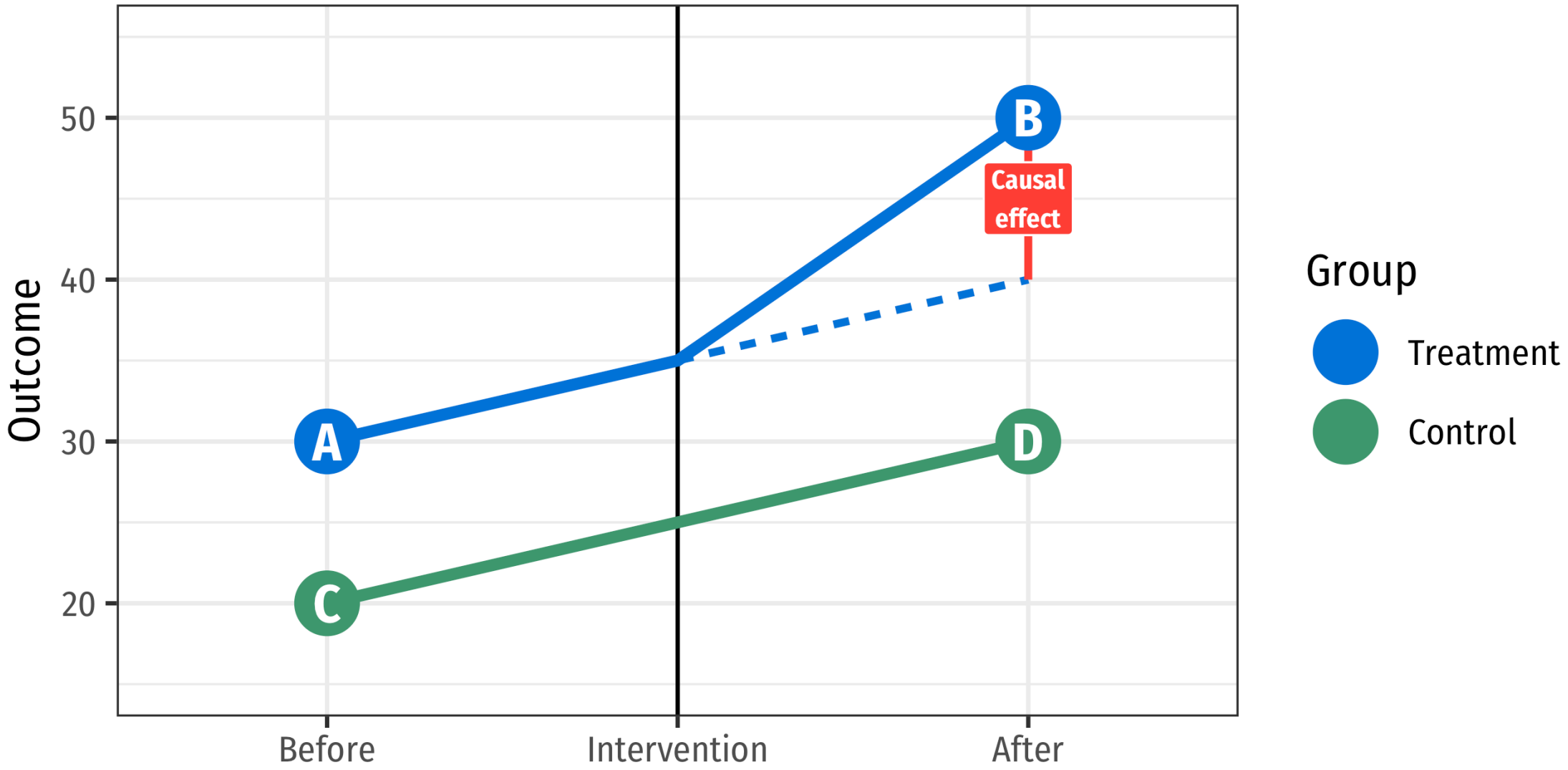
**Δ (treatment – control) = across-group growth**

| | Pre mean | Post mean | Δ (post – pre) |
|---|---|---|---|
| **Control** | **A** (never treated) | **B** (never treated) | **B – A** |
| **Treatment** | **C** (not yet treated) | **D** (treated) | **D – C** |
| **Δ (treatment – control)** | **C – A** | **D – B** | **(D – C) – (B – A)** *or* **(D – B) – (C – A)** |

**$\Delta_{\text{within units}} - \Delta_{\text{within groups}} =$ Difference-in-differences = causal effect!**

$$\begin{aligned}\text{DD} = &\left(\bar{x}_{\text{treatment, post}} - \bar{x}_{\text{treatment, pre}}\right) \\ &- \left(\bar{x}_{\text{control, post}} - \bar{x}_{\text{control, pre}}\right)\end{aligned}$$

| | Pre mean | Post mean | Δ (post − pre) |
|---|---|---|---|
| **Pennsylvania** | **23.33**<br>A | **21.17**<br>B | **-2.16**<br>B − A |
| **New Jersey** | **20.44**<br>C | **21.03**<br>D | **0.59**<br>D − C |
| **Δ (NJ − PA)** | **-2.89**<br>C − A | **-0.14**<br>D − B | **(0.59) − (−2.16) =**<br>**2.76** |

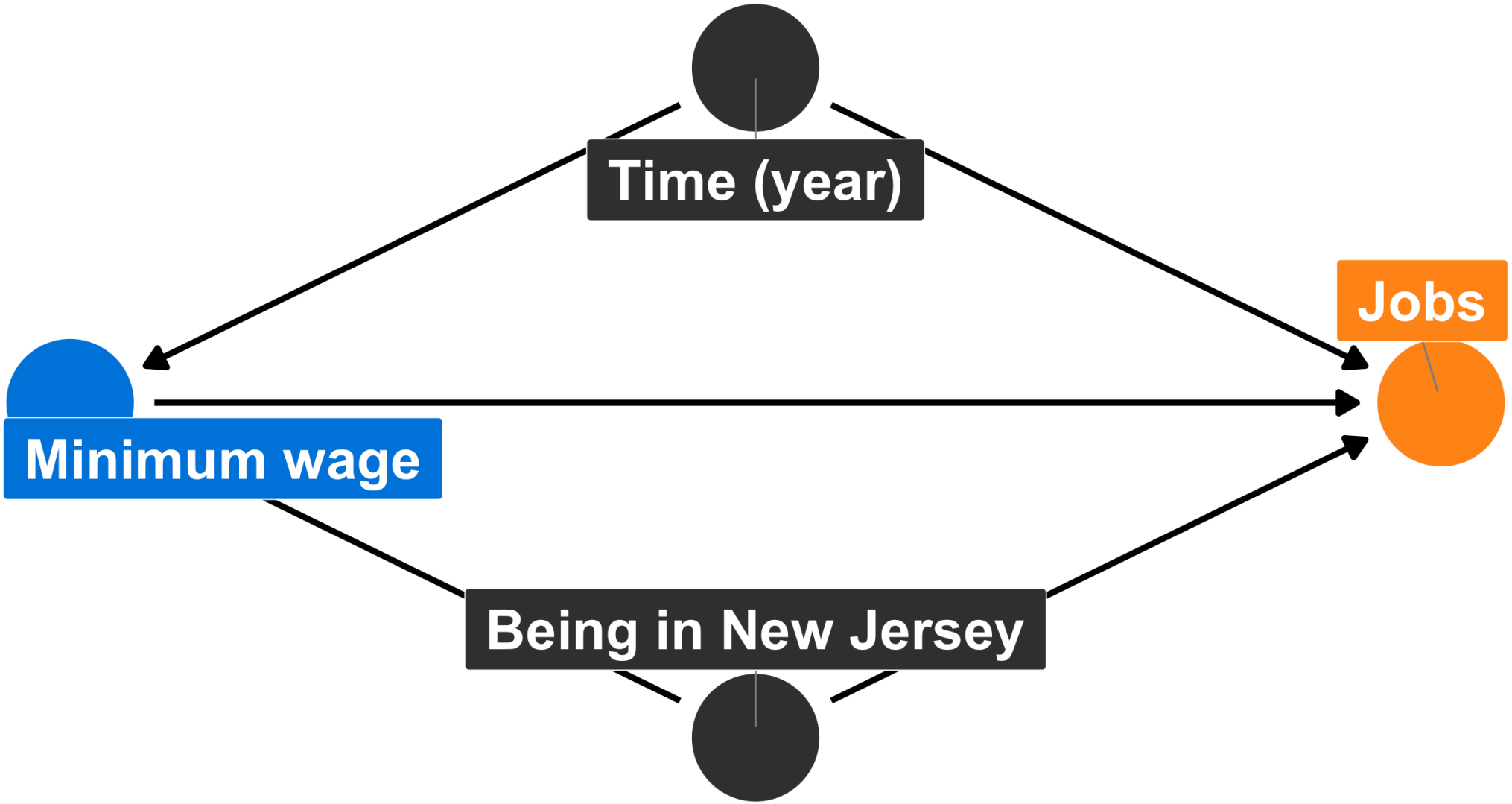

Outcome
50
40
30
20
A
B
C
D
Causal effect
Before
Intervention
After
Group
Treatment
Control

# An easier way?

Finding all the group means is tedious!

What if there are other backdoors to worry about?

**Regression to the rescue!**


Time (year)
Jobs
Minimum wage
Being in New Jersey

$$Y_{it} = \alpha + \beta\,\text{Group}_i + \gamma\,\text{Time}_t + \delta\,(\text{Group}_i \times \text{Time}_t) + \varepsilon_{it}$$

**Group = 1 or TRUE if treatment**

**Time = 1 or TRUE if after**

$$Y_{it} = \alpha + \beta\,\text{Group}_i + \gamma\,\text{Time}_t + \delta\,(\text{Group}_i \times \text{Time}_t) + \varepsilon_{it}$$

**α = Mean of control, pre-treatment**

**β = Increase in outcome across groups**

**γ = Increase in outcome over time within units**

**δ = Difference in differences!**

$$Y_{it} = \alpha + \beta\,\text{Group}_i + \gamma\,\text{Time}_t + \delta\,(\text{Group}_i \times \text{Time}_t) + \varepsilon_{it}$$

| | Pre mean | Post mean | Δ (post – pre) |
|---|---|---|---|
| Control | α | α + γ | γ |
| Treatment | α + β | α + β + γ + δ | γ + δ |
| Δ (trtmt – ctrl) | β | β + δ | δ |

.pull-left[

.small-code[

OPEN ACCESS

# Gotta catch'em all! Pokémon GO and physical activity among young adults: difference in differences study

Katherine B Howe,[1,2] Christian Suharlim,[3] Peter Ueda,[4,5] Daniel Howe, Ichiro Kawachi,[2] Eric B Rimm[1,6,7]

[1]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, USA

[2]Department of Social and Behavioral Sciences, Harvard TH Chan School of Public Health, Boston, MA, USA

[3]Center for Health and Decision Science, Department of Health Policy and Management, Harvard TH Chan School of Public Health, Boston, MA, USA

[4]Department of Global Health and Population, Harvard TH Chan School of Public Health, Boston, MA, USA

[5]Clinical Epidemiology Unit, Department of Medicine, Solna,

## ABSTRACT

### OBJECTIVE

To estimate the effect of playing Pokémon GO on the number of steps taken daily up to six weeks after installation of the game.

### DESIGN

Cohort study using online survey data.

### PARTICIPANTS

Survey participants of Amazon Mechanical Turk (n=1182) residing in the United States, aged 18 to 35 years and using iPhone 6 series smartphones.

### MAIN OUTCOME MEASURES

Number of daily steps taken each of the four weeks before and six weeks after installation of Pokémon

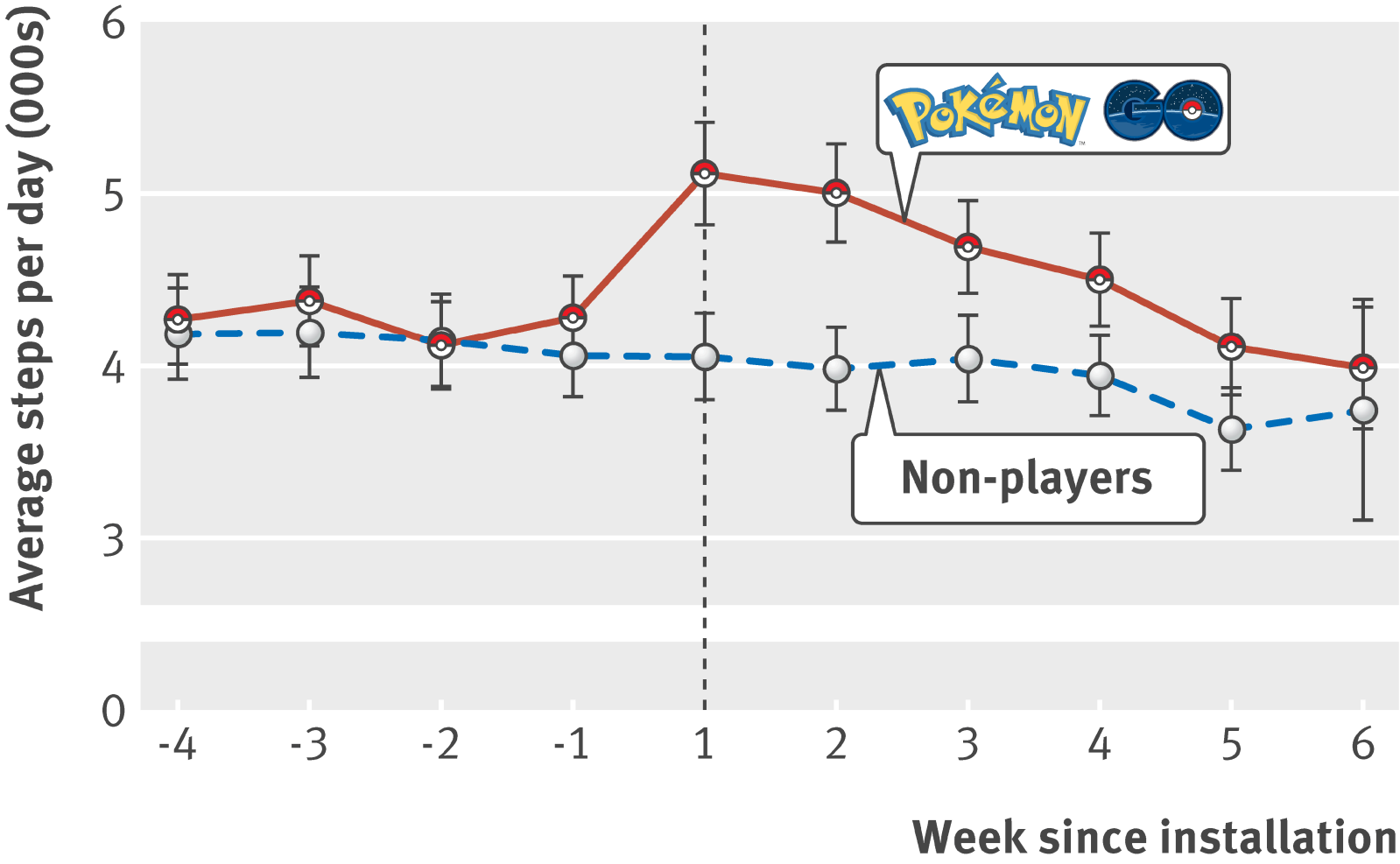
### CONCLUSIONS

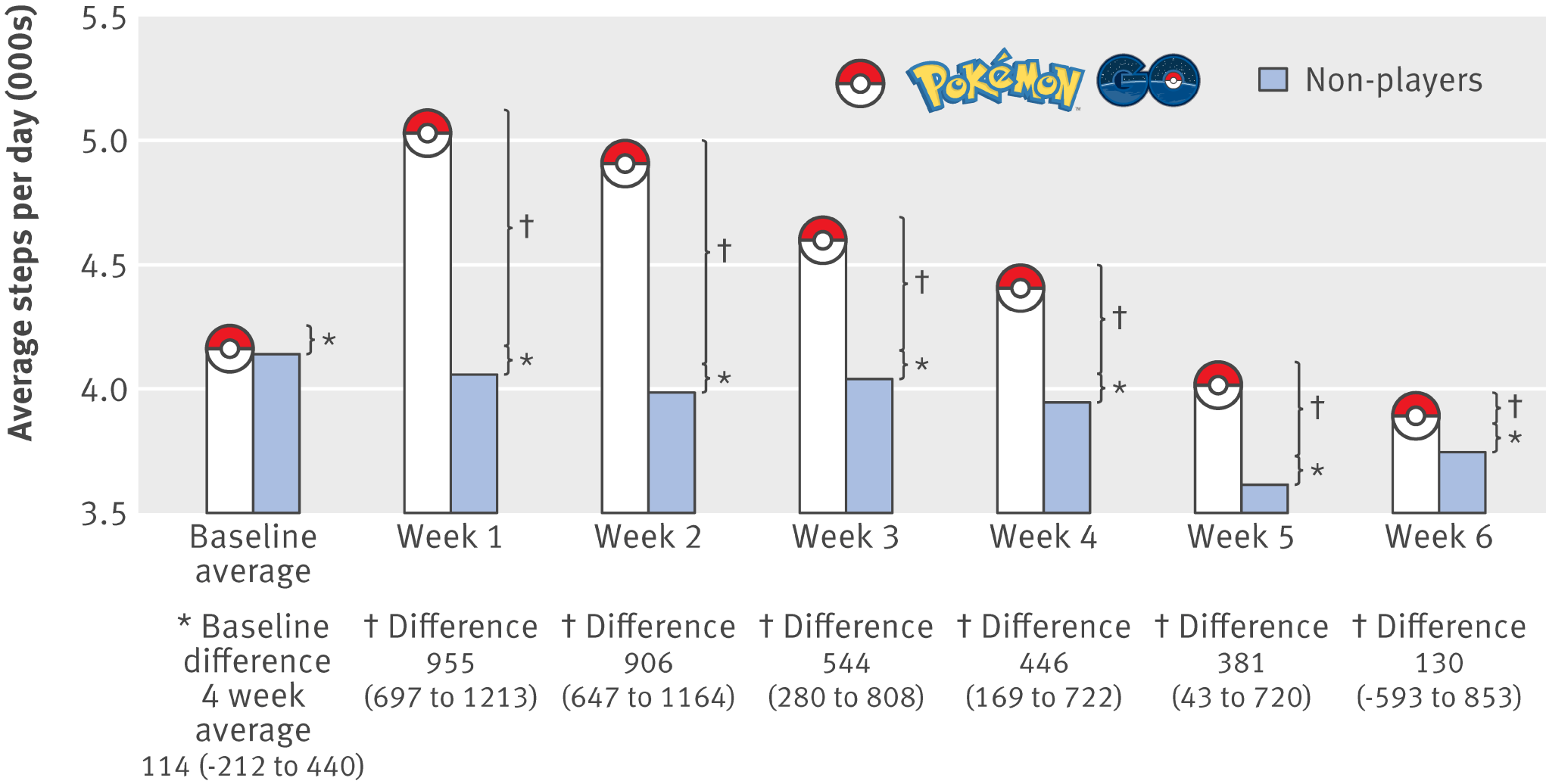
Pokémon GO was associated with an increase in the daily number of steps after installation of the game. The association was, however, moderate and no longer observed after six weeks.

## Introduction

Pokémon GO is an augmented reality game in which players search real world locations for cartoon characters appearing on their smartphone screen. Since its launch in July 2016, the game has been downloaded over 500 million times worldwide.

Games that incentivise exercise might have the potential to promote and sustain physical activity habits.[1,2] Because walking is encouraged while playing,


Average steps per day (000s)
6
5
4
3
0
-4
-3
-2
-1
1
2
3
4
5
6
Week since installation
Pokémon GO
Non-players

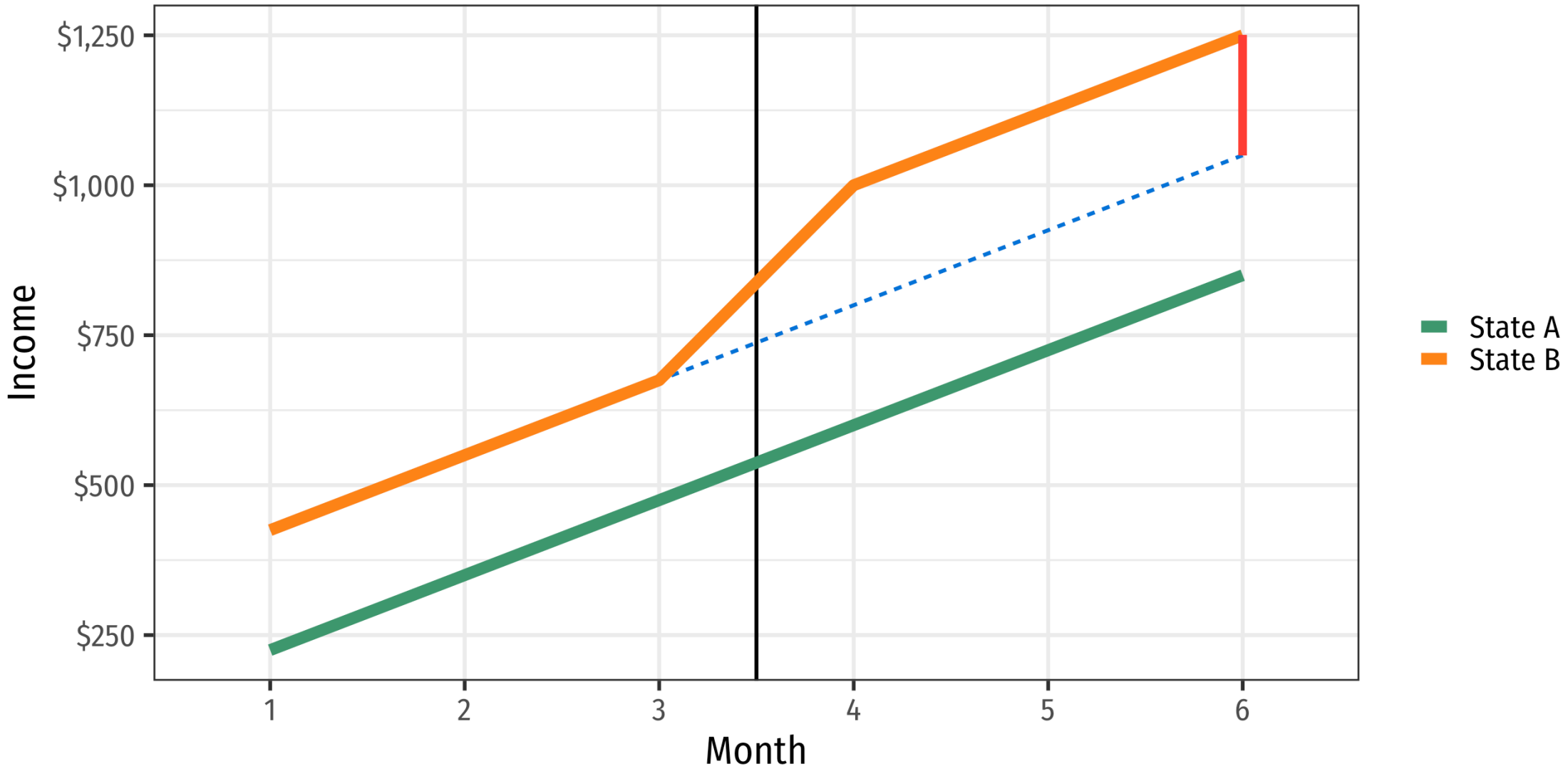

Average steps per day (000s)
5.5
5.0
4.5
4.0
3.5
Non-players
Baseline average
Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
* Baseline difference 4 week average
114 (-212 to 440)
† Difference
955
(697 to 1213)
† Difference
906
(647 to 1164)
† Difference
544
(280 to 808)
† Difference
446
(169 to 722)
† Difference
381
(43 to 720)
† Difference
130
(-593 to 853)

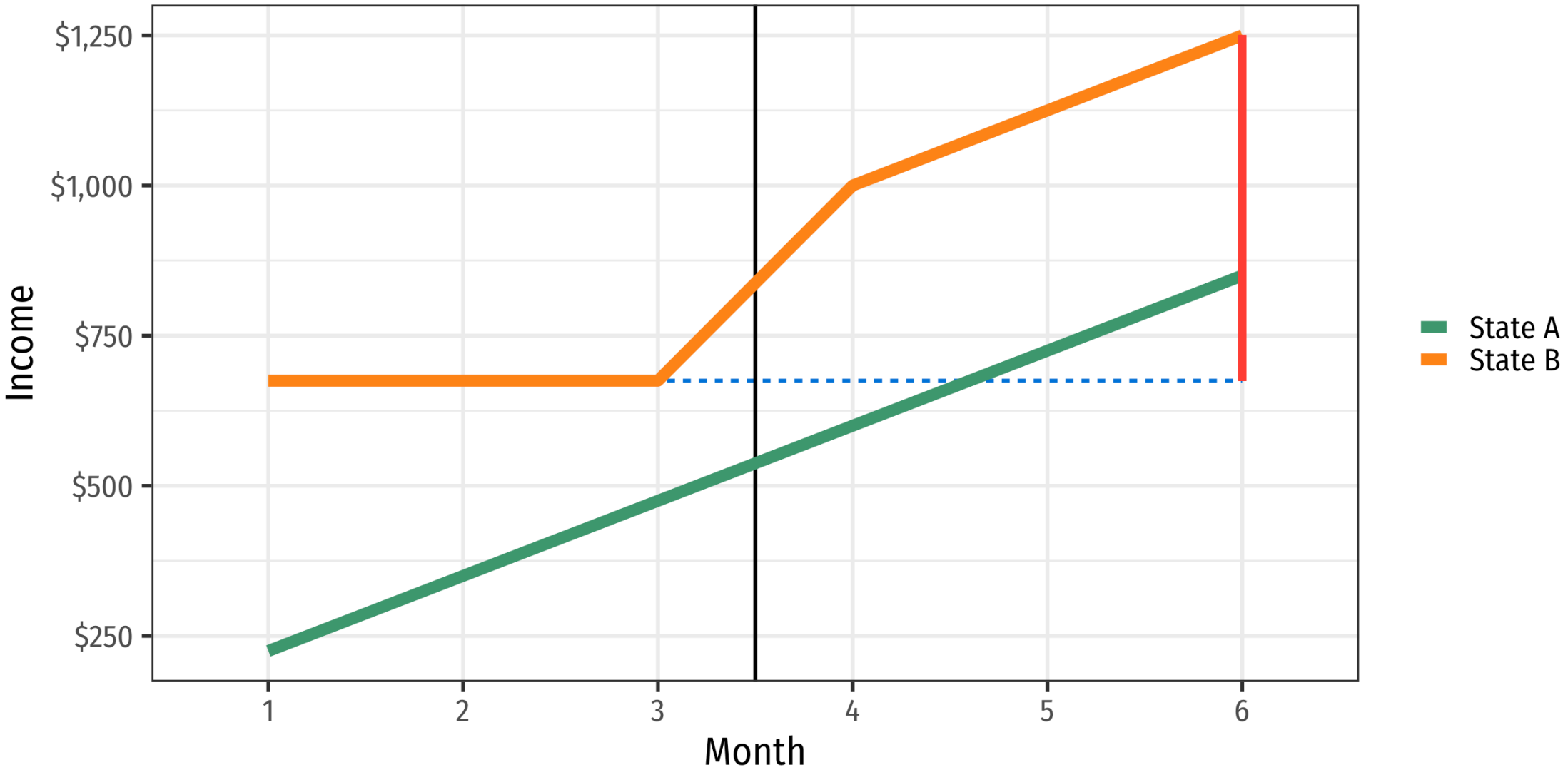
# Diff-in-diff assumptions

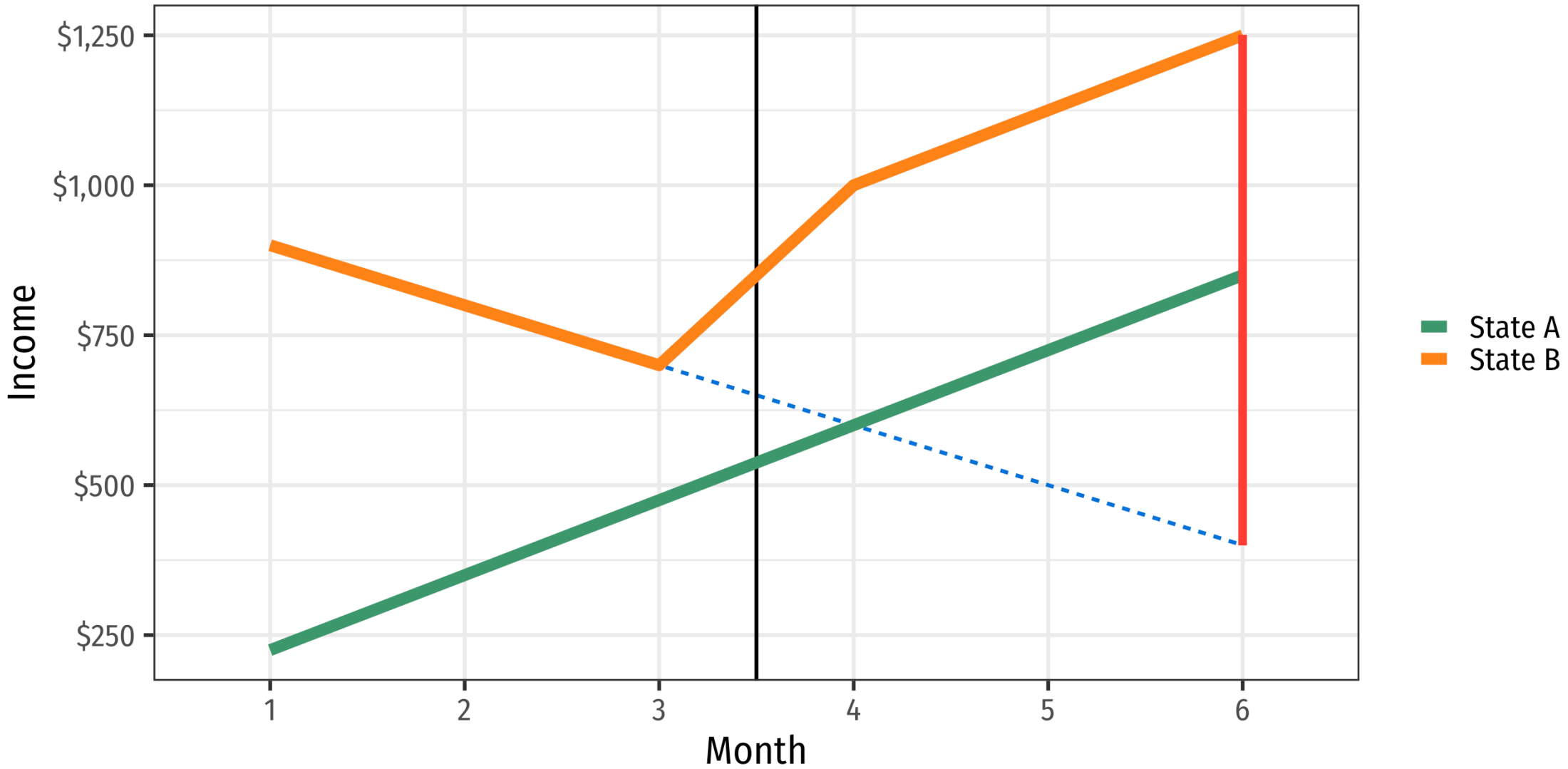
# Assumptions

## Parallel trends

**Treatment and control groups might have different values at first, but we assume that the treatment group would have changed like the control group in the absence of treatment**


$1,250
$1,000
$750
$500
$250
Income
1
2
3
4
5
6
Month
State A
State B

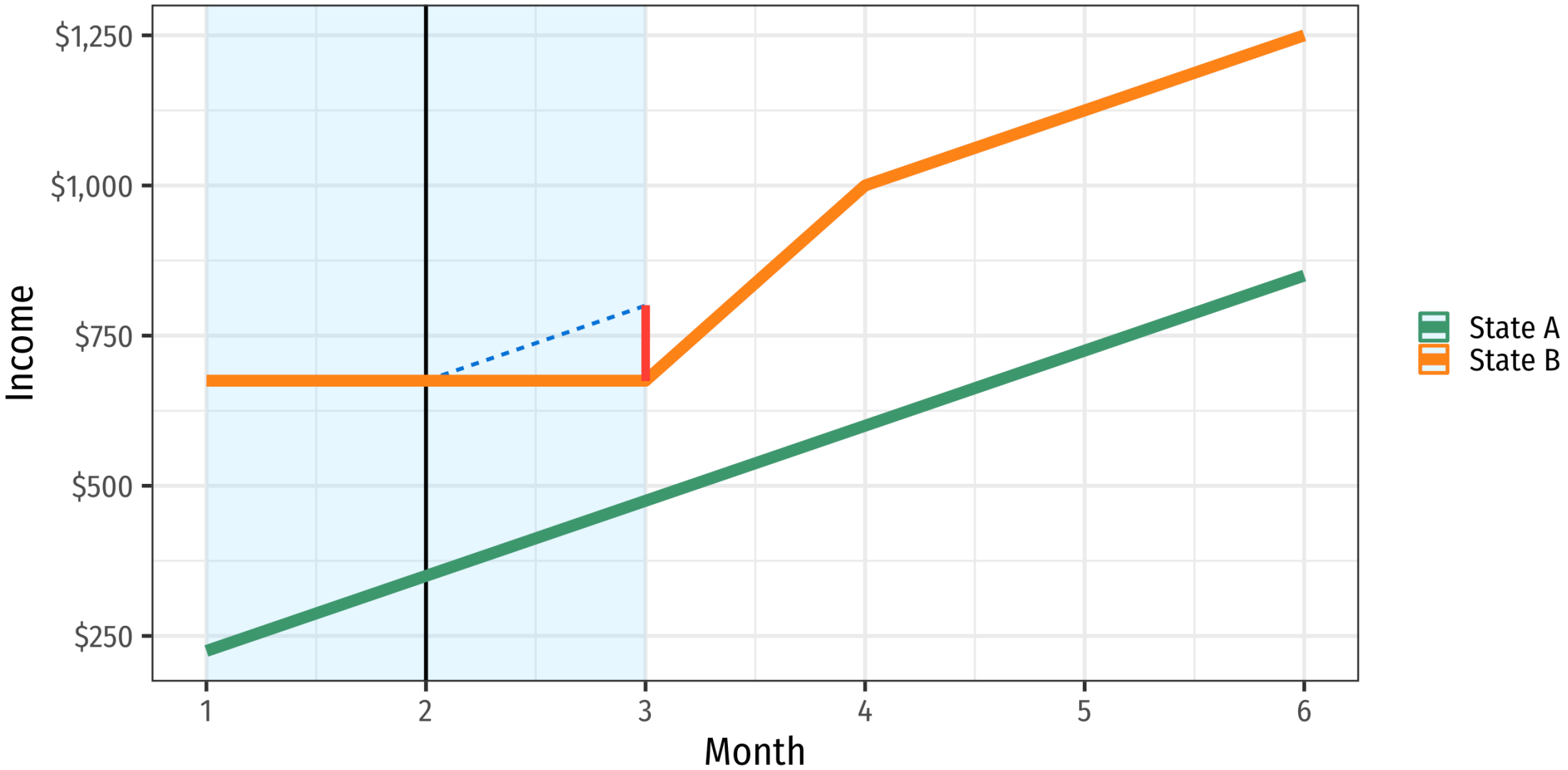

$1,250
$1,000
$750
$500
$250
Income
1
2
3
4
5
6
Month
State A
State B

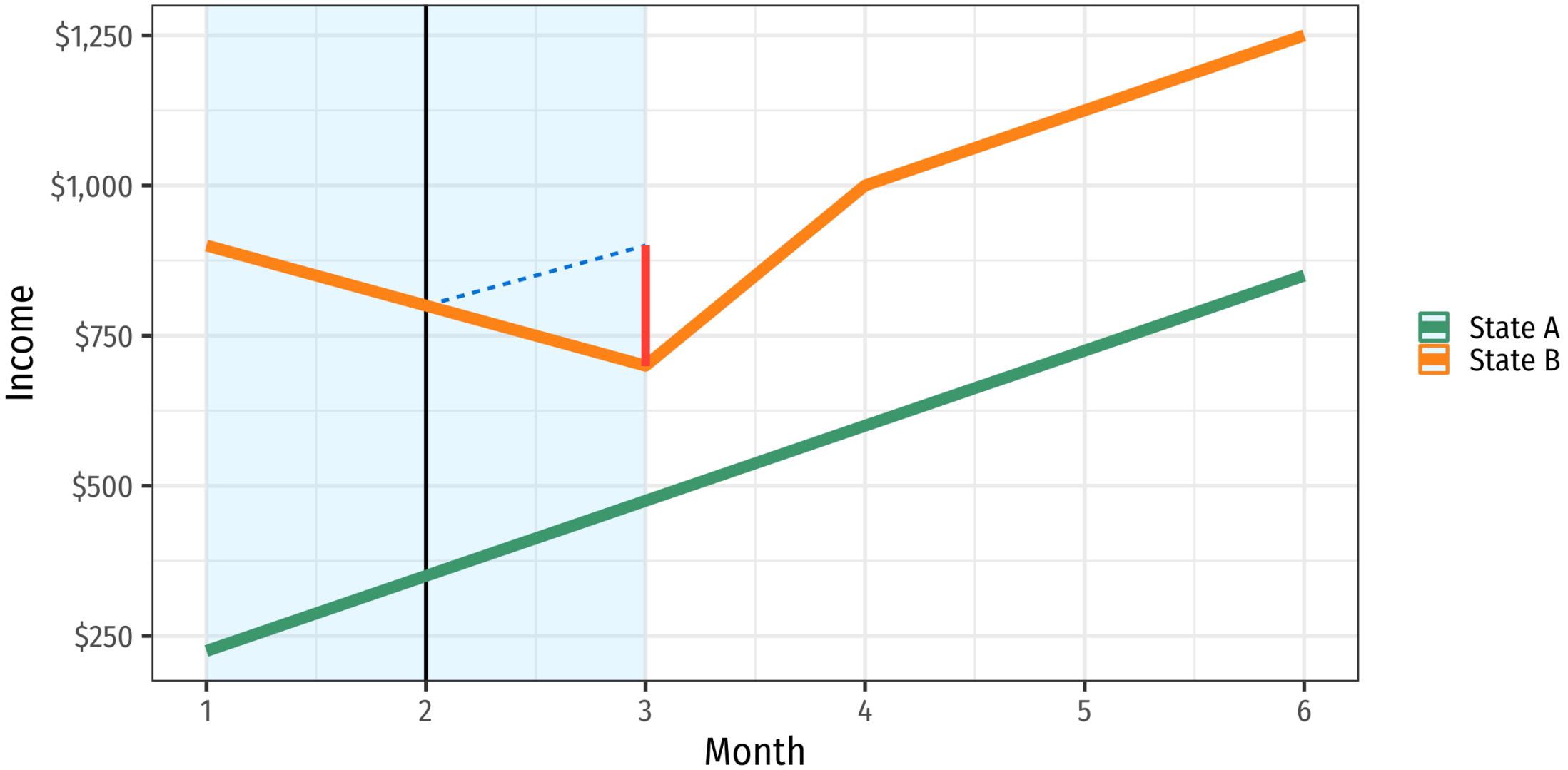

$1,250
$1,000
$750
$500
$250
Income
1
2
3
4
5
6
Month
State A
State B

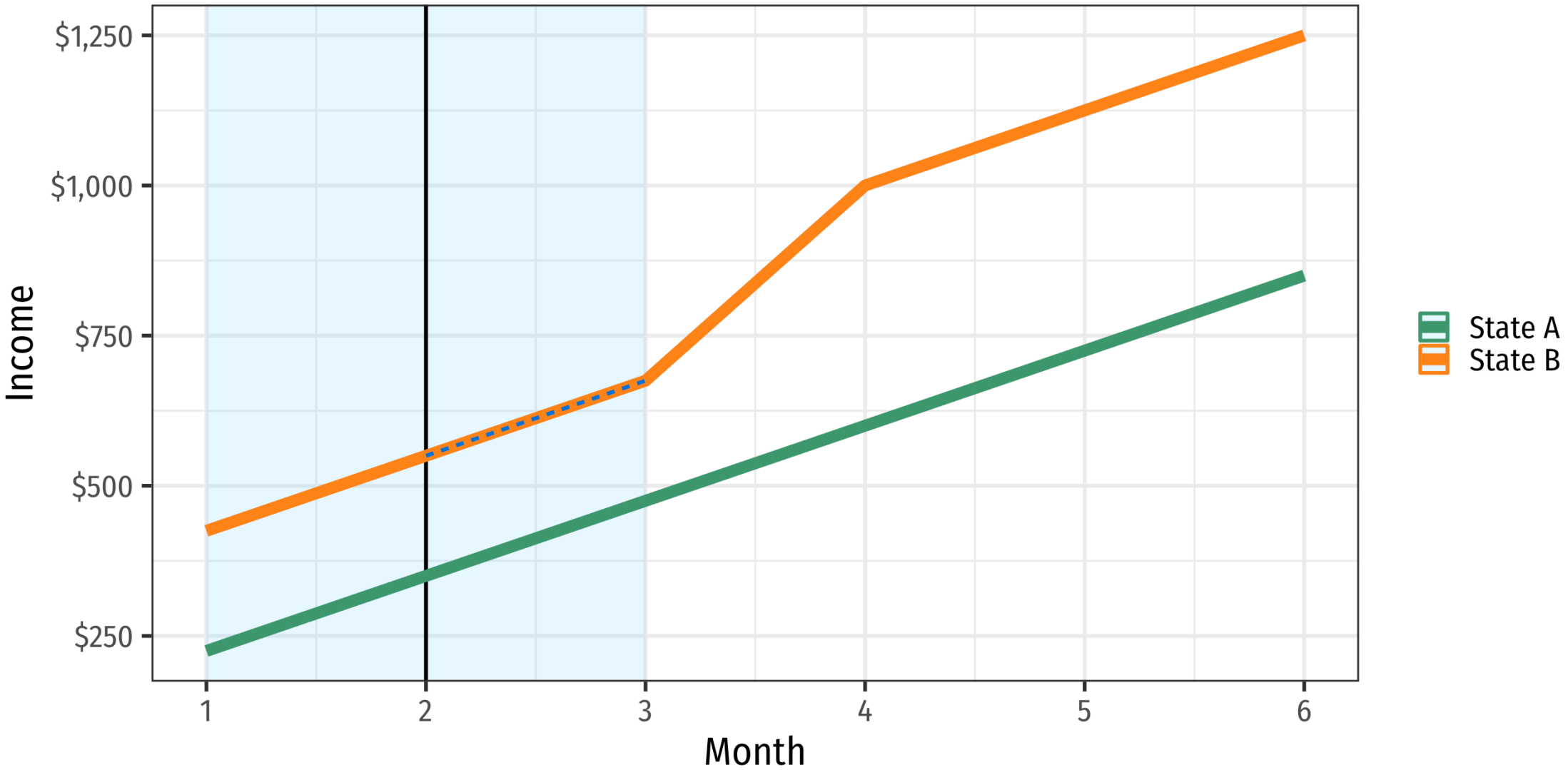
# Assumptions

## Parallel trends

**Check by pretending the treatment happened earlier; if there's an effect, there's likely an underlying trend**


$1,250
$1,000
$750
$500
$250
Income
1
2
3
4
5
6
Month
State A
State B

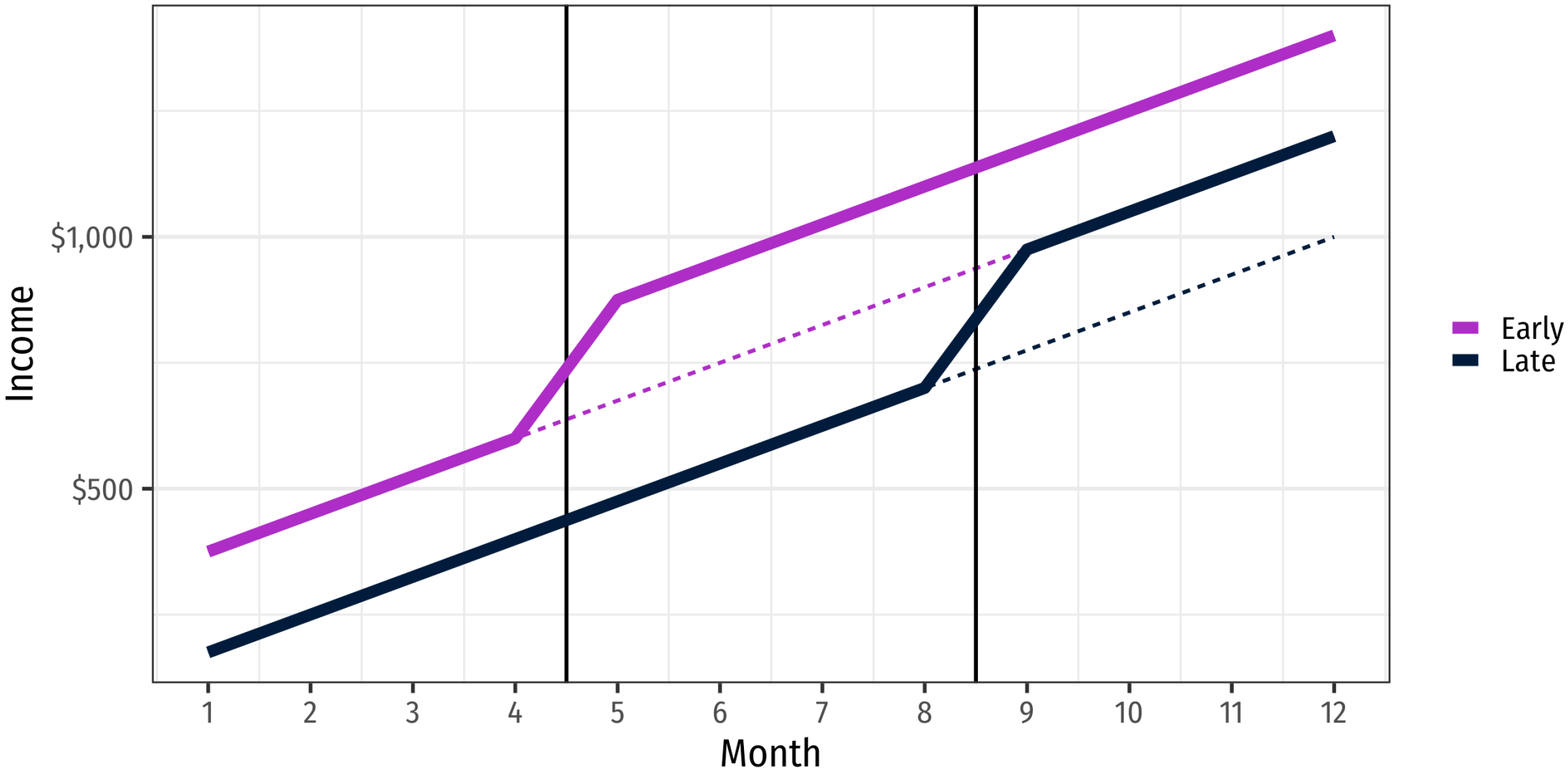

$1,250
$1,000
$750
$500
$250
Income
1
2
3
4
5
6
Month
State A
State B

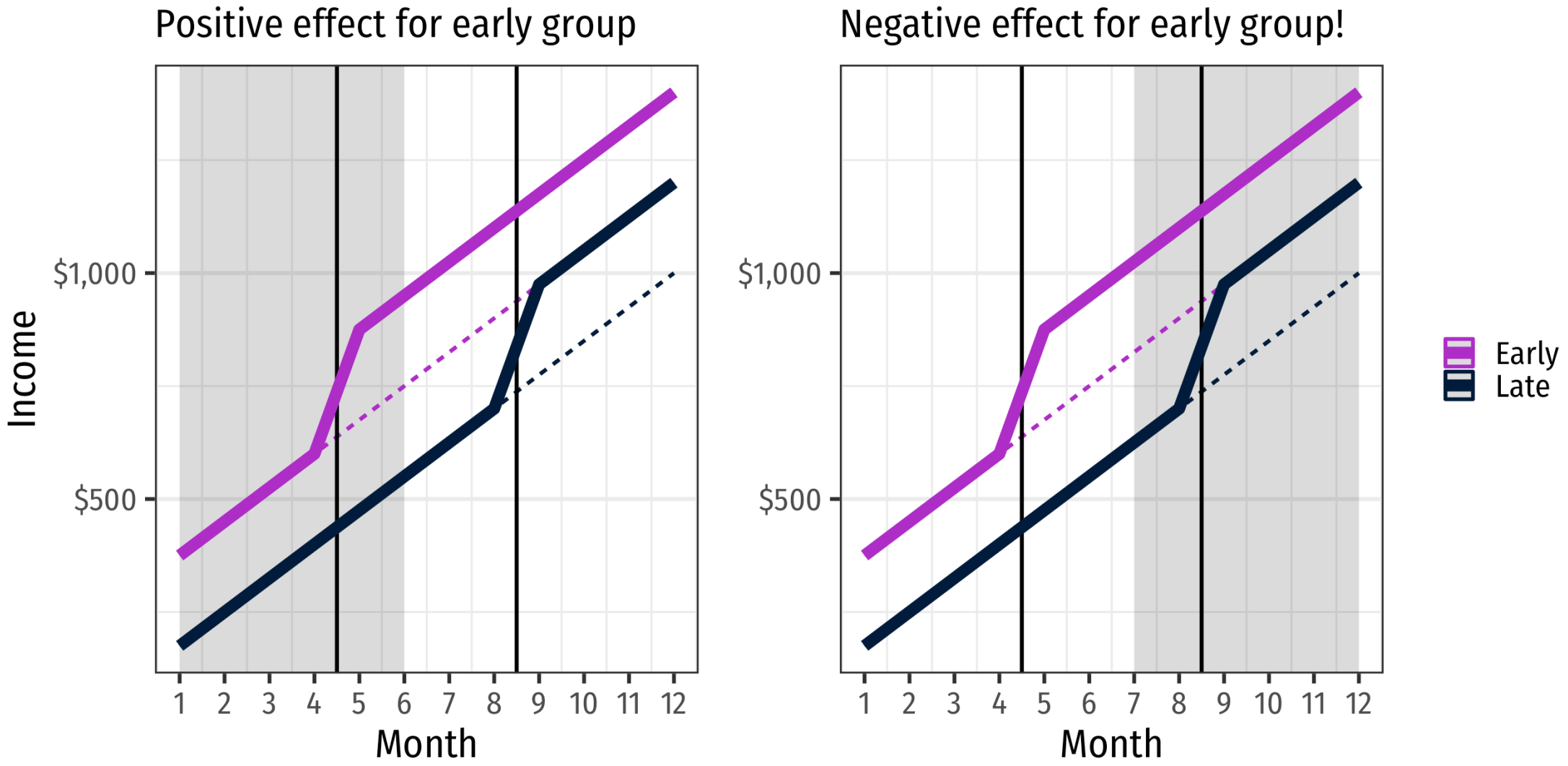

$1,250
$1,000
$750
$500
$250
Income
1
2
3
4
5
6
Month
State A
State B

# Assumptions

## Treatment timing

**Units often receive treatment at different times, which can distort your estimate!**


Income
$1,000
$500
1
2
3
4
5
6
7
8
9
10
11
12
Month
Early
Late


Positive effect for early group
Negative effect for early group!
Income
$1,000
$500
1 2 3 4 5 6 7 8 9 10 11 12
Month
Early
Late

# Assumptions

## You can check how big of an issue this is with Goodman-Bacon decomposition

## R package: `bacondecomp`

**DIFFERENCE-IN-DIFFERENCES WITH VARIATION IN TREATMENT TIMING***

Andrew Goodman-Bacon

July 2019

Abstract: The canonical difference-in-differences (DD) estimator contains two time periods, "pre" and "post", and two groups, "treatment" and "control". Most DD applications, however, exploit variation across groups of units that receive treatment at different times. This paper shows that the general estimator equals a weighted average of all possible two-group/two-period DD estimators in the data. This defines the DD estimand and identifying assumption, a generalization of common trends. I discuss how to interpret DD estimates and propose a new balance test. I show how to decompose the difference between two specifications, and provide a new analysis of models that include time-varying controls.