ITSx: Policy Analysis Using Interrupted Time Series

Week 1 Slides

Michael Law, Ph.D.
The University of British Columbia

COURSE OVERVIEW

Layout of the weeks

1. Introduction, setup, data sources
2. Single series interrupted time series analysis
3. ITS with a control group
4. Regression discontinuities and extensions
5. Course wrap-up

Overview of Week 1

- Overview of observational study designs
  - Threats to validity
  - Major common study designs
- Introduction to ITS
  - Steps in conducting a study
  - Discuss time periods, cohorts, and outcomes
- Optional Introduction to R and RStudio

THREATS TO VALIDITY

Internal and External Validity

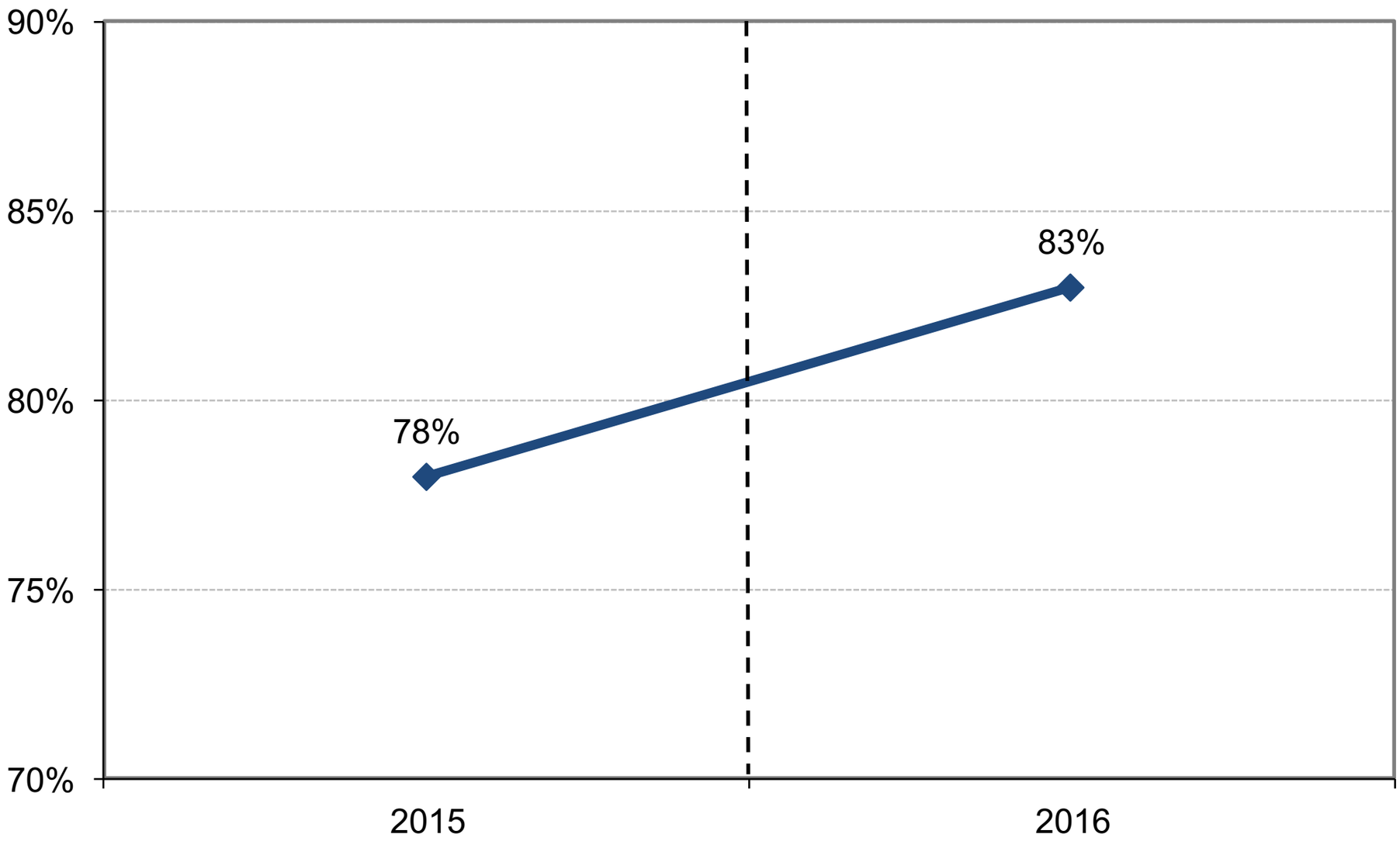
- **Internal validity:** the validity of the conclusions within the study population

- **External validity:** how the results of your (internally valid) study generalize to the rest of the world

Motivating Example

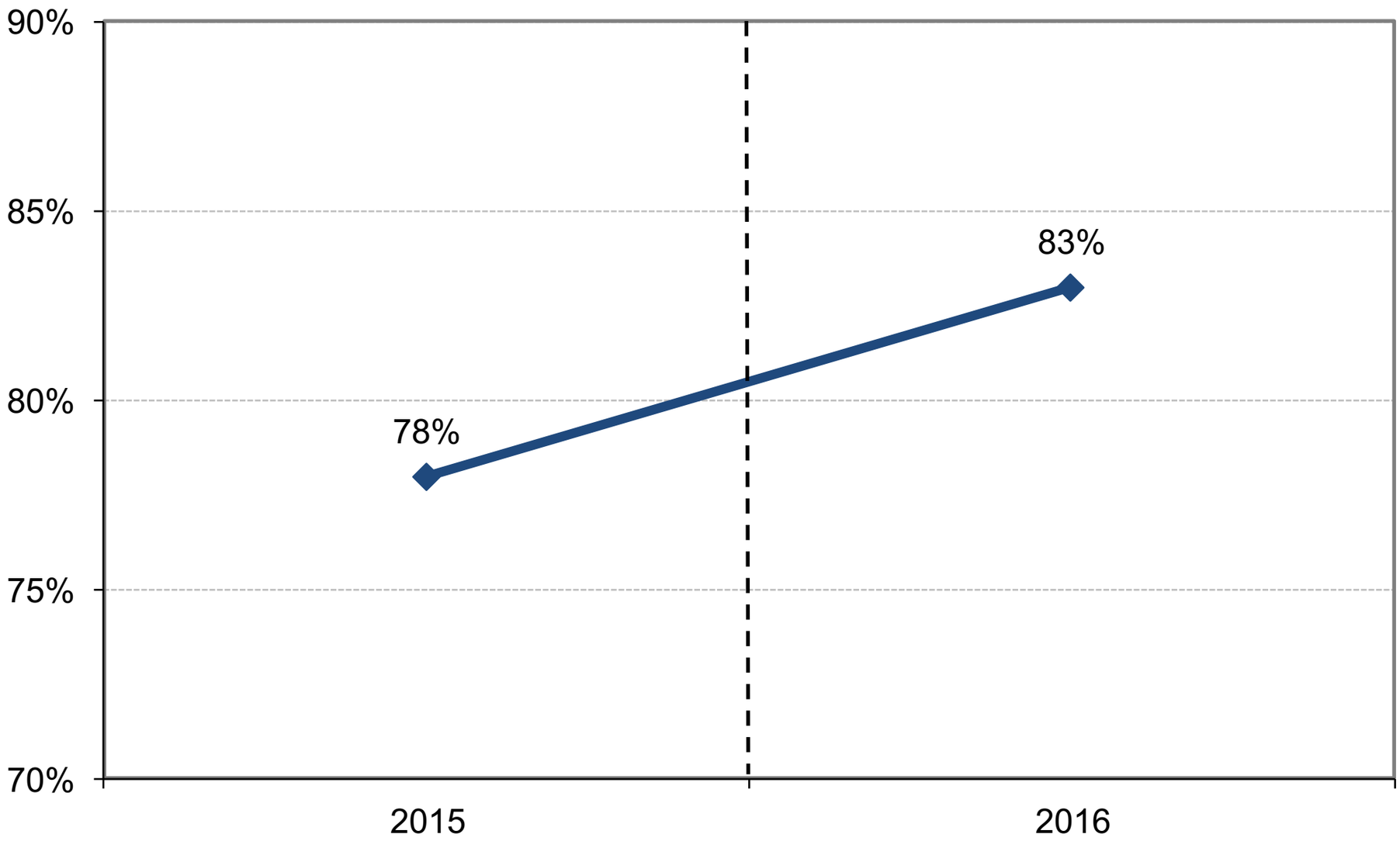
The average grade in ITSx in 2015 was 78%. In 2016, edX capped class sizes to encourage more interaction with the the instructor. In that year, the average grade was 83%.

Question: What effect did changing class sizes have on student performance?


90%
85%
80%
75%
70%
78%
83%
2015
2016

The Counterfactual

- The outcome that you would have observed absent the intervention being studied
- The problem: it's unobservable
- The "solution": estimate it


90%
85%
80%
75%
70%
78%
83%
2015
2016

THREATS TO INTERNAL VALIDITY

Threats to Internal Validity

Five major threats to validity:

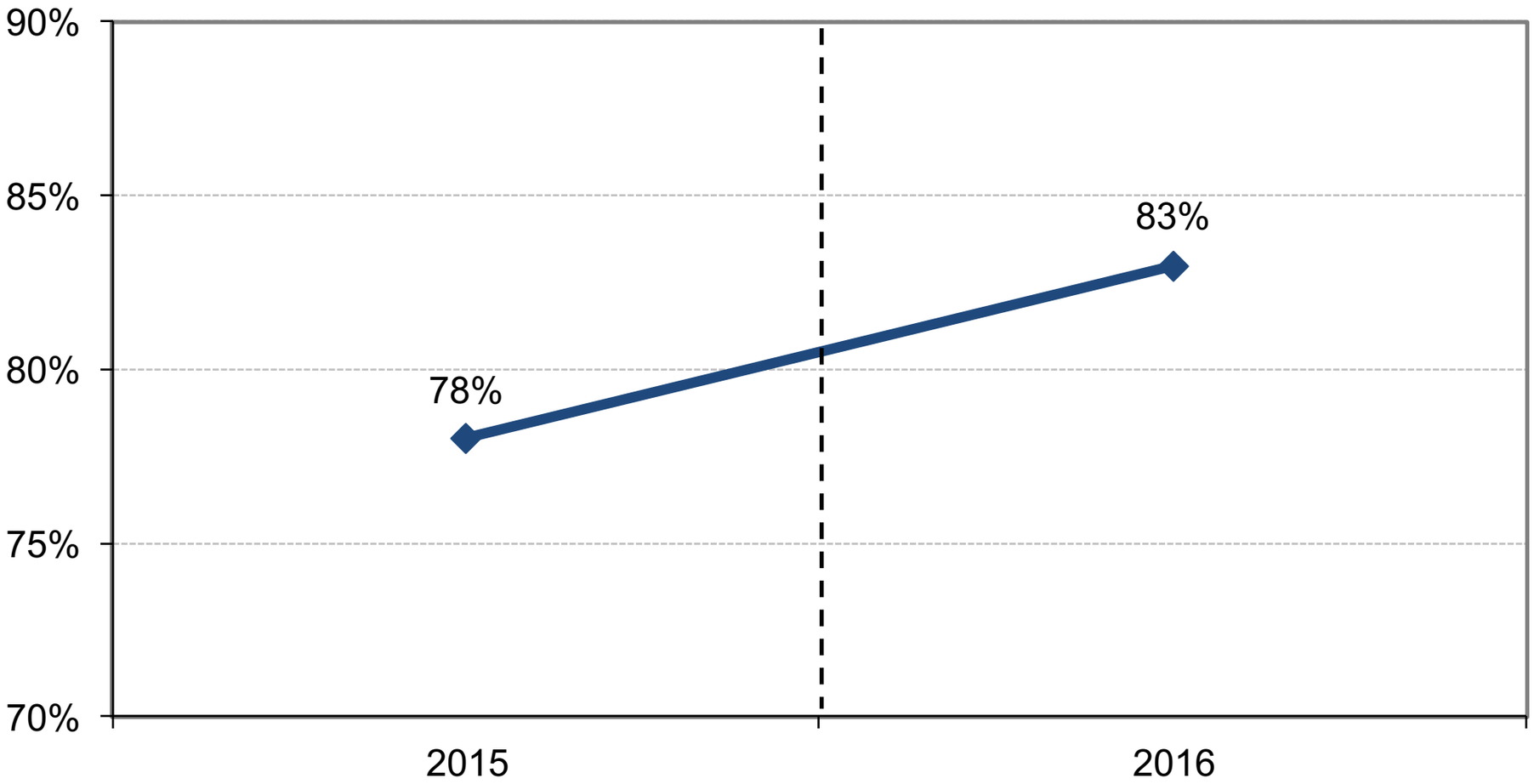
1. History
2. Maturation
3. Instrumentation & Testing
4. Statistical Regression
5. Selection

History

- An event occurring between pre- and post-measurement that is not the intervention of interest

Example:

In 2016, edX also started a scholarship program that waived the course fee for top students.

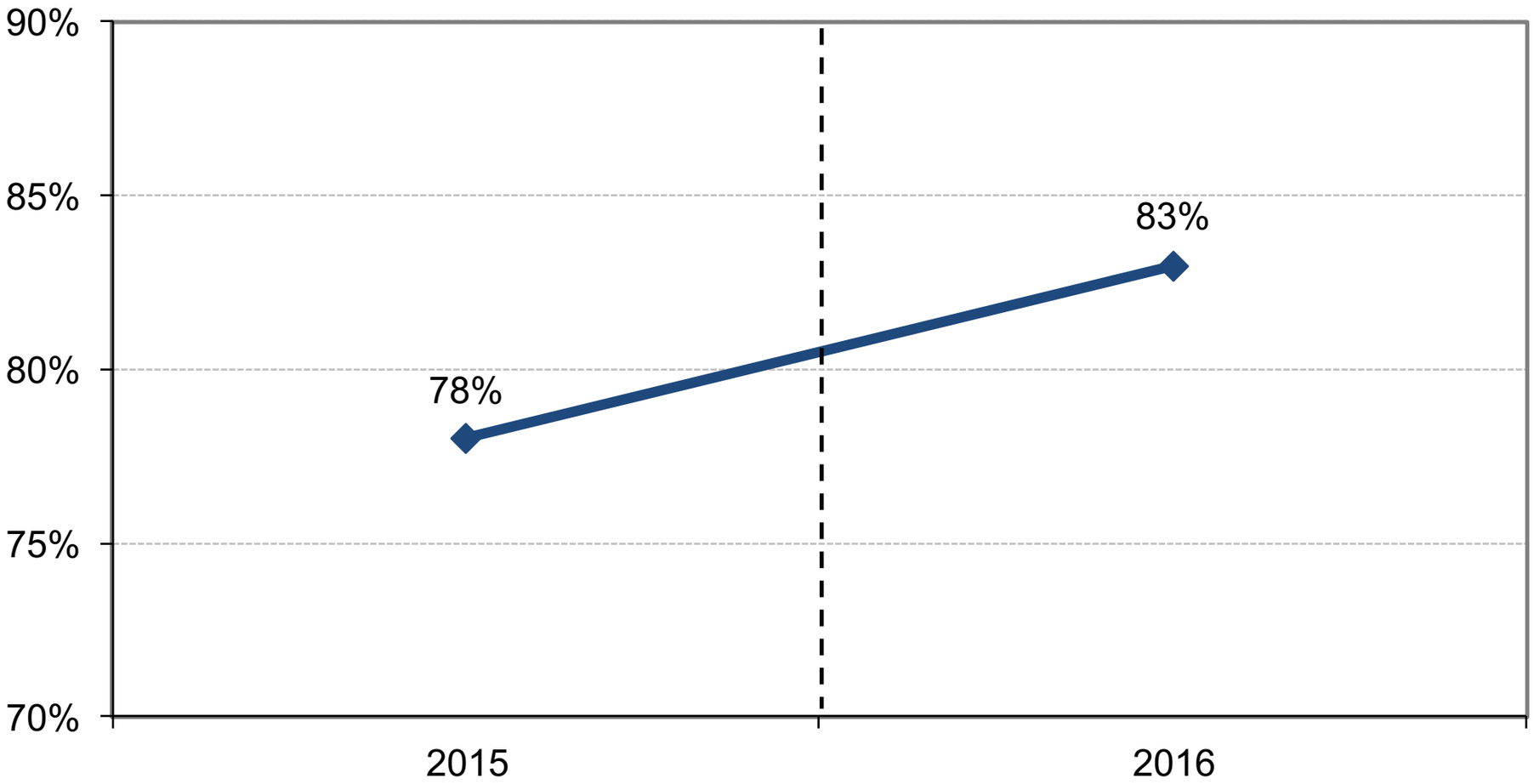

90%
85%
80%
75%
70%
78%
83%
2015
2016

Maturation

- Subjects growing older, wiser, healthier, etc. between pre- and post- measurement

Example:

As students become more accustomed to learning online, they become better at mastering the material.


90%
85%
80%
75%
70%
78%
83%
2015
2016

Instrumentation & Testing

- A change in outcome measurement, either:
  - A change in the measure itself, or
  - familiarity with the test by the participants

Example:

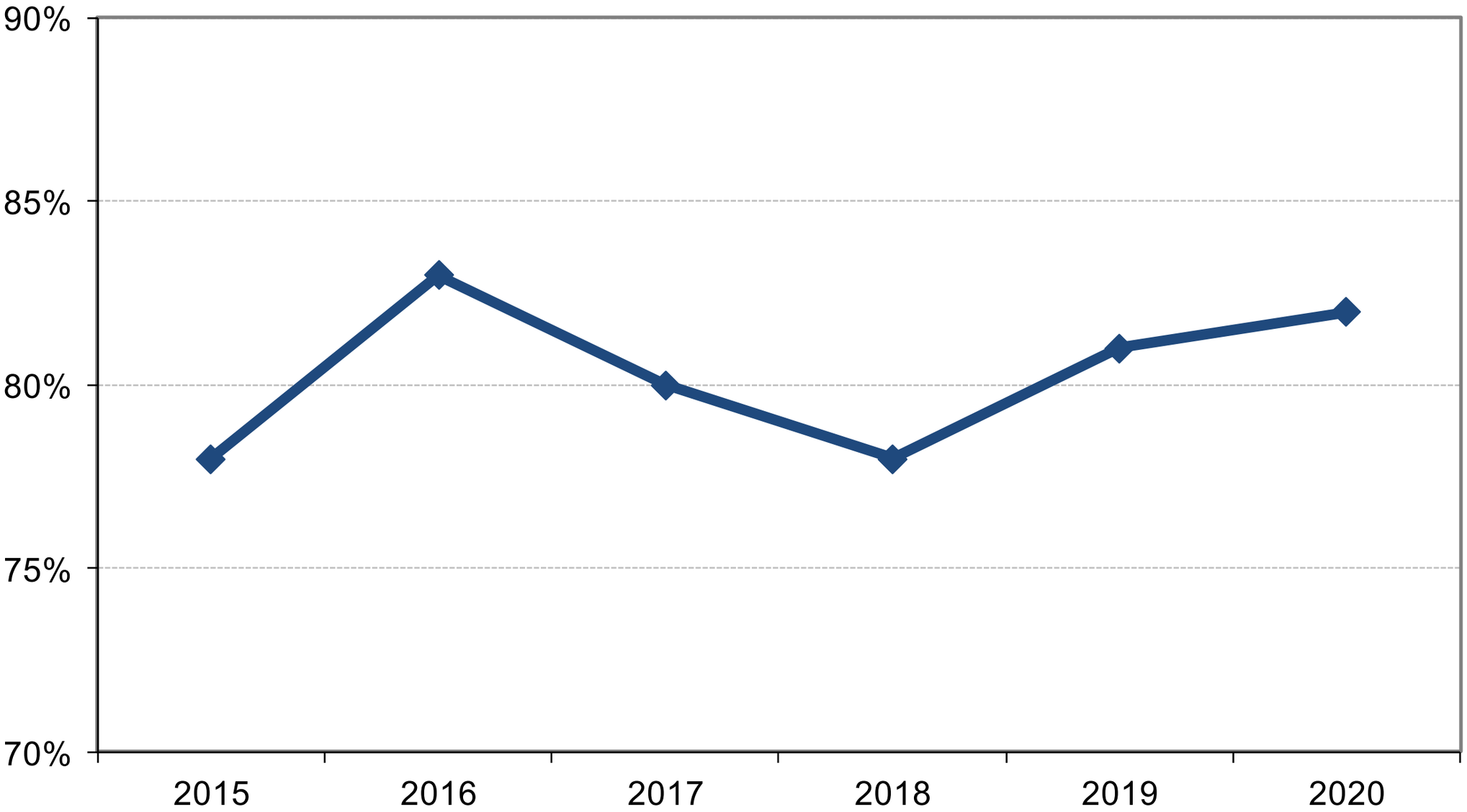
I make the edX tests easier when class sizes become smaller because I know all of you more personally

Statistical Regression

- A value which naturally varies over time changes to a different value

Example:

The quality of students randomly varies by year. In the year after the change, ITSx attracts stronger students by chance.


90%
85%
80%
75%
70%
2015
2016
2017
2018
2019
2020

Selection

- Pre-intervention differences between people in one group vs. another

Example:

Students who take a first course offering differ from those who take a “wait and see” approach in ways that are not measurable.

Interactions

When more than one group is compared, threats can interact with one another

- Selection-Maturation
- Selection-Instrumentation
- Selection-History
- Selection-Regression

OBSERVATIONAL STUDY DESIGNS

What is an Observational Study

- Where the researcher does not control the intervention or factor being studied
- Selection mechanism can be potentially problematic, especially if it's choice-based
- Confounders and existing trends likely differ between groups

Describing study design

- Notation: X to denote interventions, O to denote observations, and dashed lines to show non-random groups

$$O_1 \quad X \quad O_2$$

1. **Post-only with control:**

$$\begin{array}{ll} X & O_1 \\ \hline & O_1 \end{array}$$

2. **Pre-post:** $O_1 \quad X \quad O_2$

3. **Pre-post with control:**

$$\begin{array}{lll} O_1 & X & O_2 \\ \hline O_1 & & O_2 \end{array}$$

4. Interrupted Time Series:

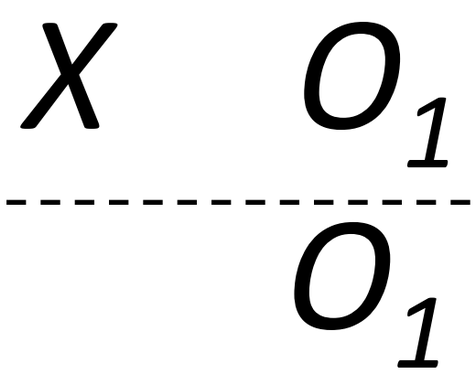
$O_1\ O_2\ O_3\ O_4\ O_5\ O_6\ O_7\ O_8\ X\ O_9\ O_{10}\ O_{11}\ O_{12}\ O_{13}\ O_{14}\ O_{15}\ O_{16}$

5. Interrupted Time Series with Control:

$O_1\ O_2\ O_3\ O_4\ O_5\ O_6\ O_7\ O_8\ X\ O_9\ O_{10}\ O_{11}\ O_{12}\ O_{13}\ O_{14}\ O_{15}\ O_{16}$

$O_1\ O_2\ O_3\ O_4\ O_5\ O_6\ O_7\ O_8\ \ O_9\ O_{10}\ O_{11}\ O_{12}\ O_{13}\ O_{14}\ O_{15}\ O_{16}$

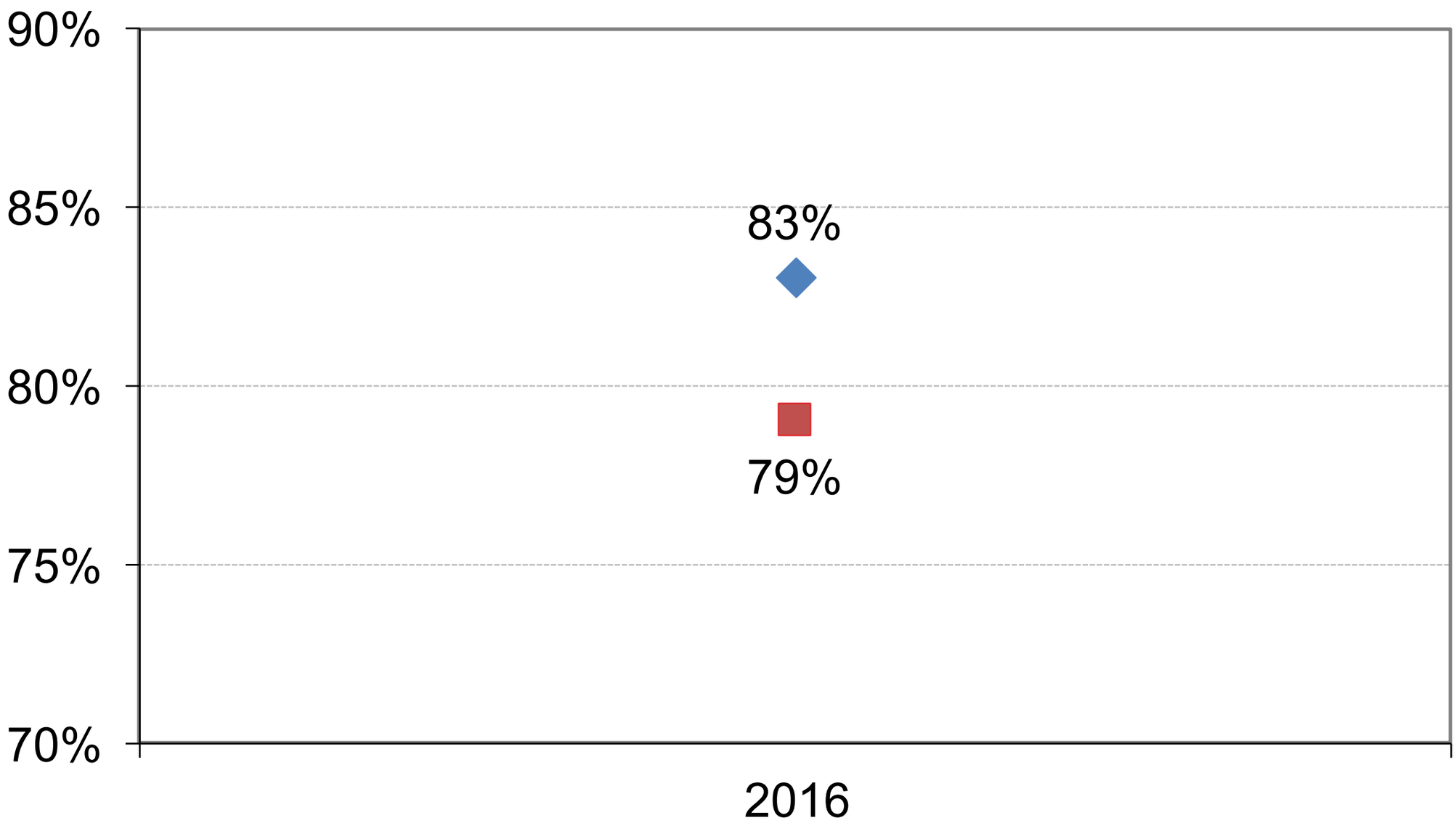
Post-only with Control



Post-only with Control

- Design
  - Compare two groups after the policy change
  - Only one group was exposed to the policy

- Counterfactual assumption
  - Groups would have had identical outcomes absent the intervention


90%
85%
80%
75%
70%
83%
79%
2016

Potential Post-only Biases

1. Selection
2. Instrumentation
3. Statistical Regression

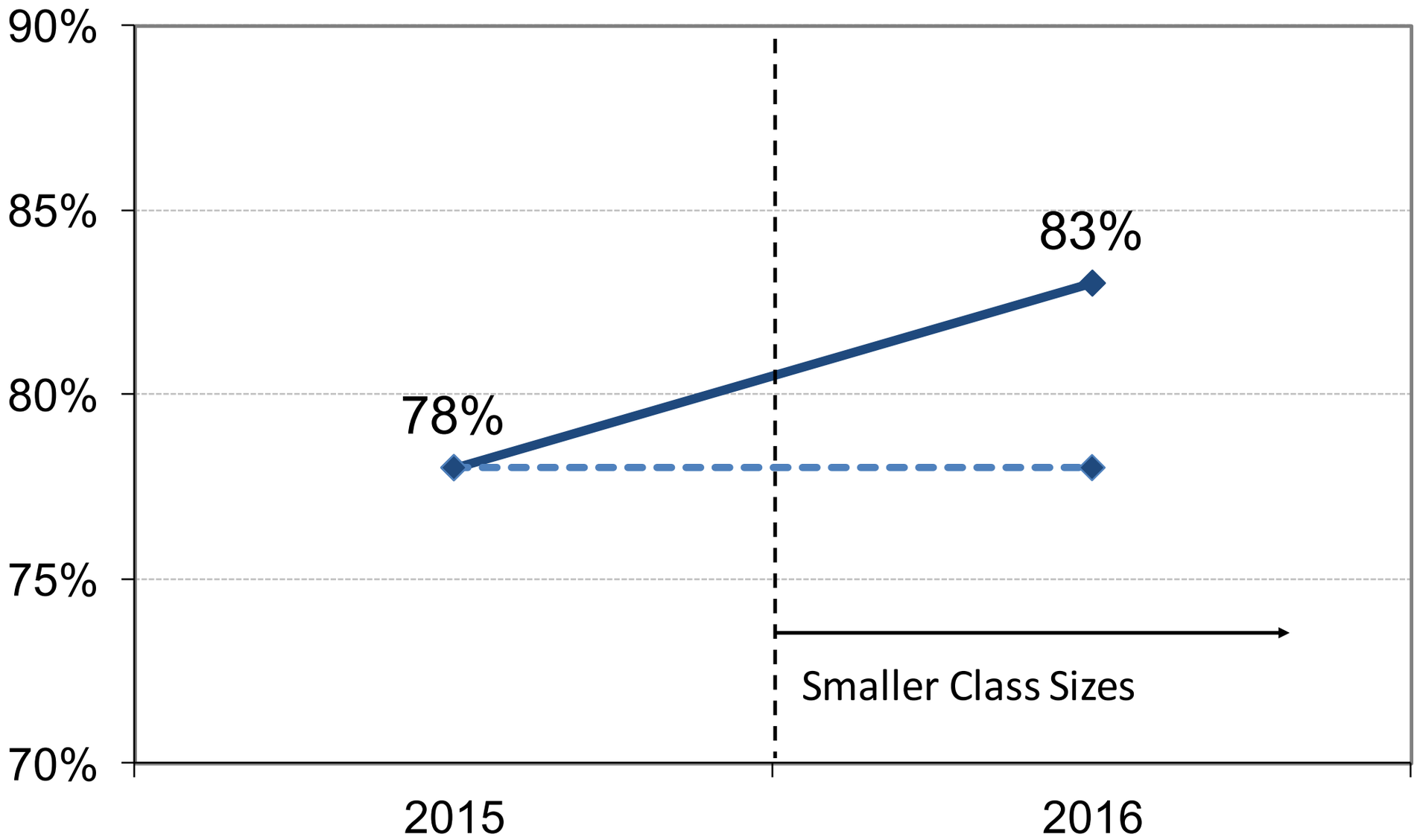
Pre-post

$$O_1 \quad X \quad O_2$$

Pre-Post Research Design

- Design
  - Compare the same group before and after the policy change

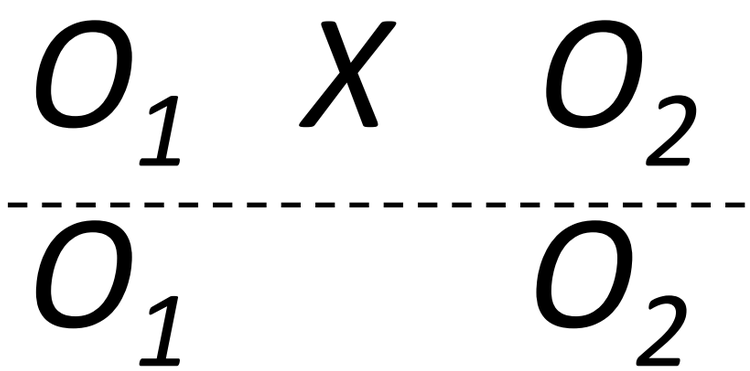
- Major Assumption
  - The outcome would not have changed absent the policy


90%
85%
80%
75%
70%
83%
78%
Smaller Class Sizes
2015
2016

Potential Pre-Post Biases

1. History
2. Maturation
3. Instrumentation
4. Statistical Regression

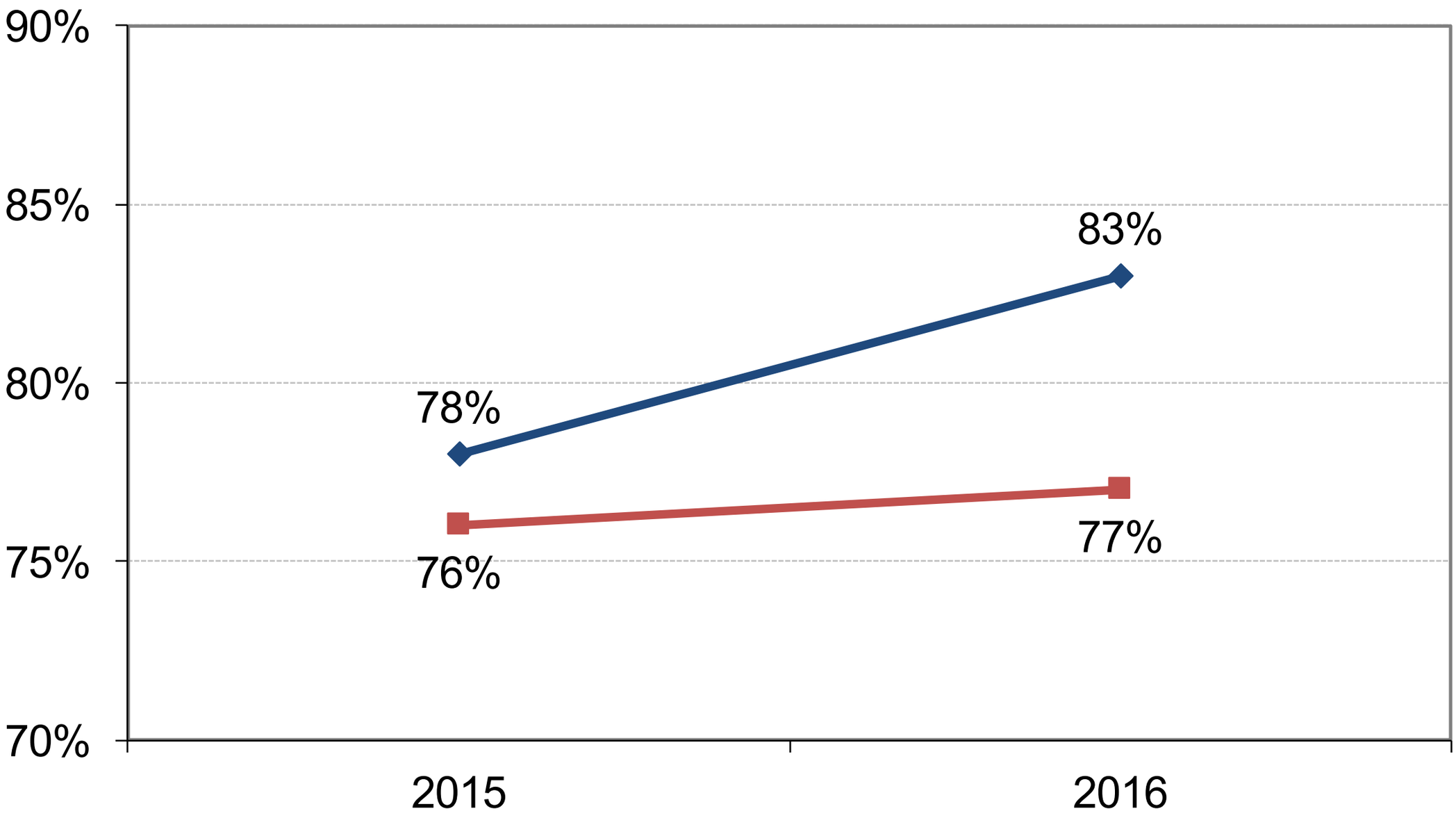
Pre-post with Control



Pre-post with Control

- Design
  - Compare two groups before and after the policy change, only one of which was exposed to the policy

- Major Assumption
  - The change in the outcome among those exposed to the policy would have been the same as those not exposed to the policy


90%
85%
80%
75%
70%
83%
78%
77%
76%
2015
2016

Potential Internal Validity Concerns

1. Selection—History
2. Selection—Maturation
3. Selection—Instrumentation

INTERRUPTED TIME SERIES

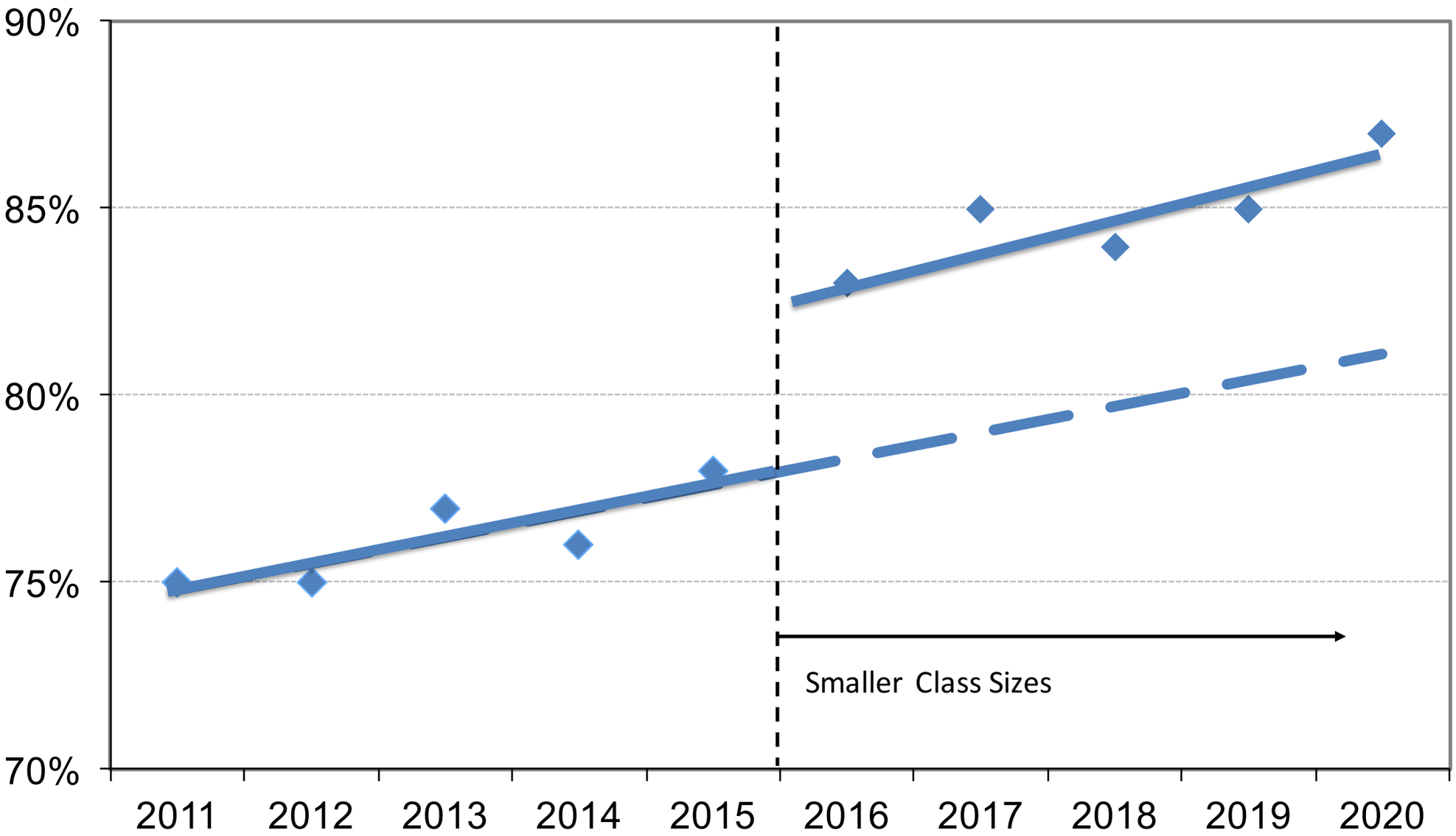
Single Group ITS

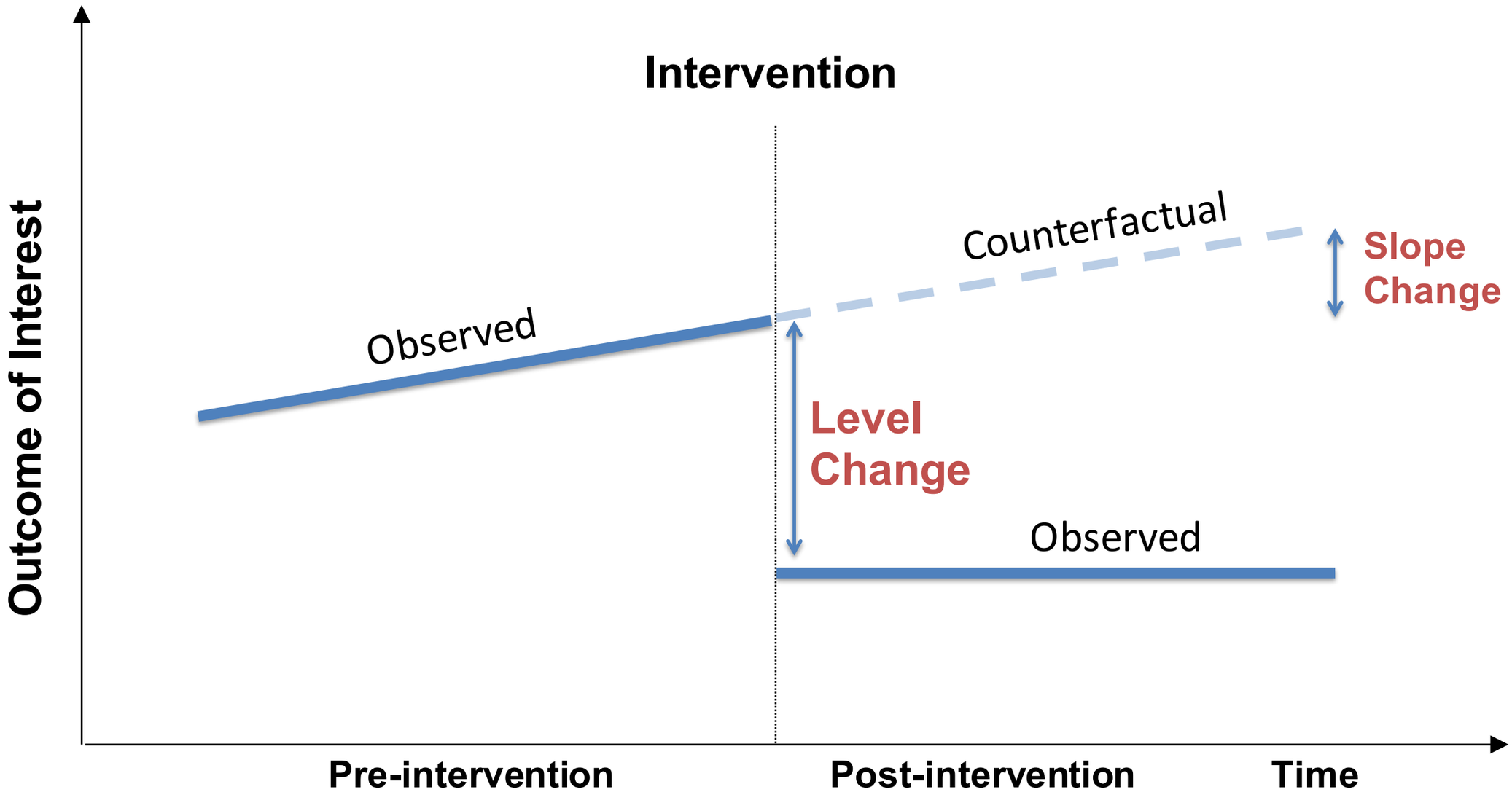
$O_1\ O_2\ O_3\ O_4\ O_5\ O_6\ O_7\ O_8\ X\ O_9\ O_{10}\ O_{11}\ O_{12}\ O_{13}\ O_{14}\ O_{15}\ O_{16}$

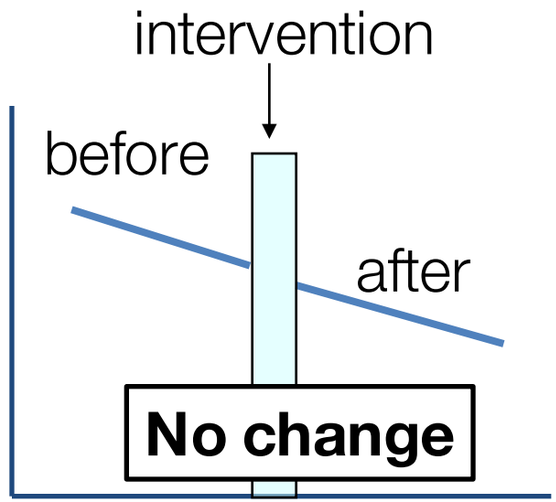
Interrupted Time Series

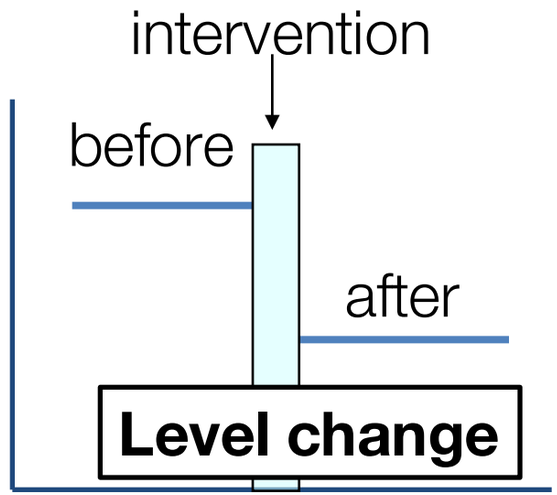
- Design
  - Compare longitudinal trends before and after the policy change

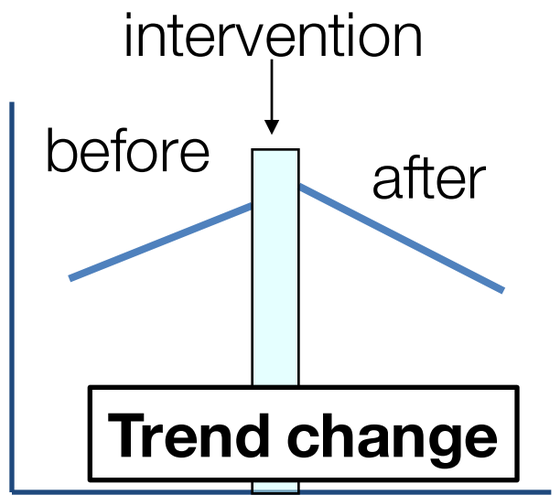
- Major Assumption
  - The existing level and trend in the outcome among those exposed to the intervention would have remained the same absent the intervention

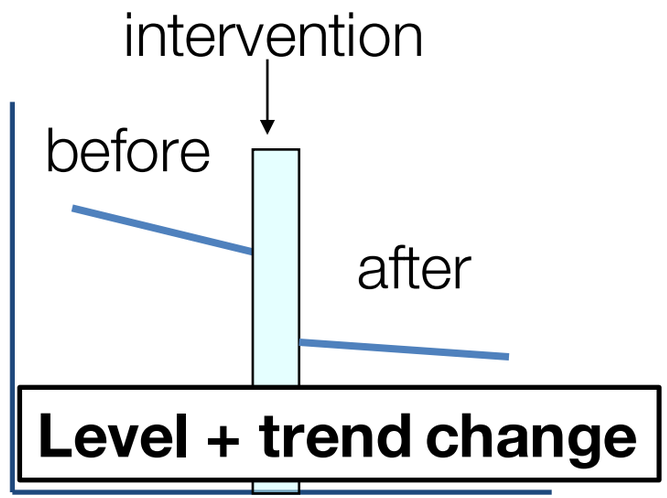

90%
85%
80%
75%
70%
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Smaller Class Sizes

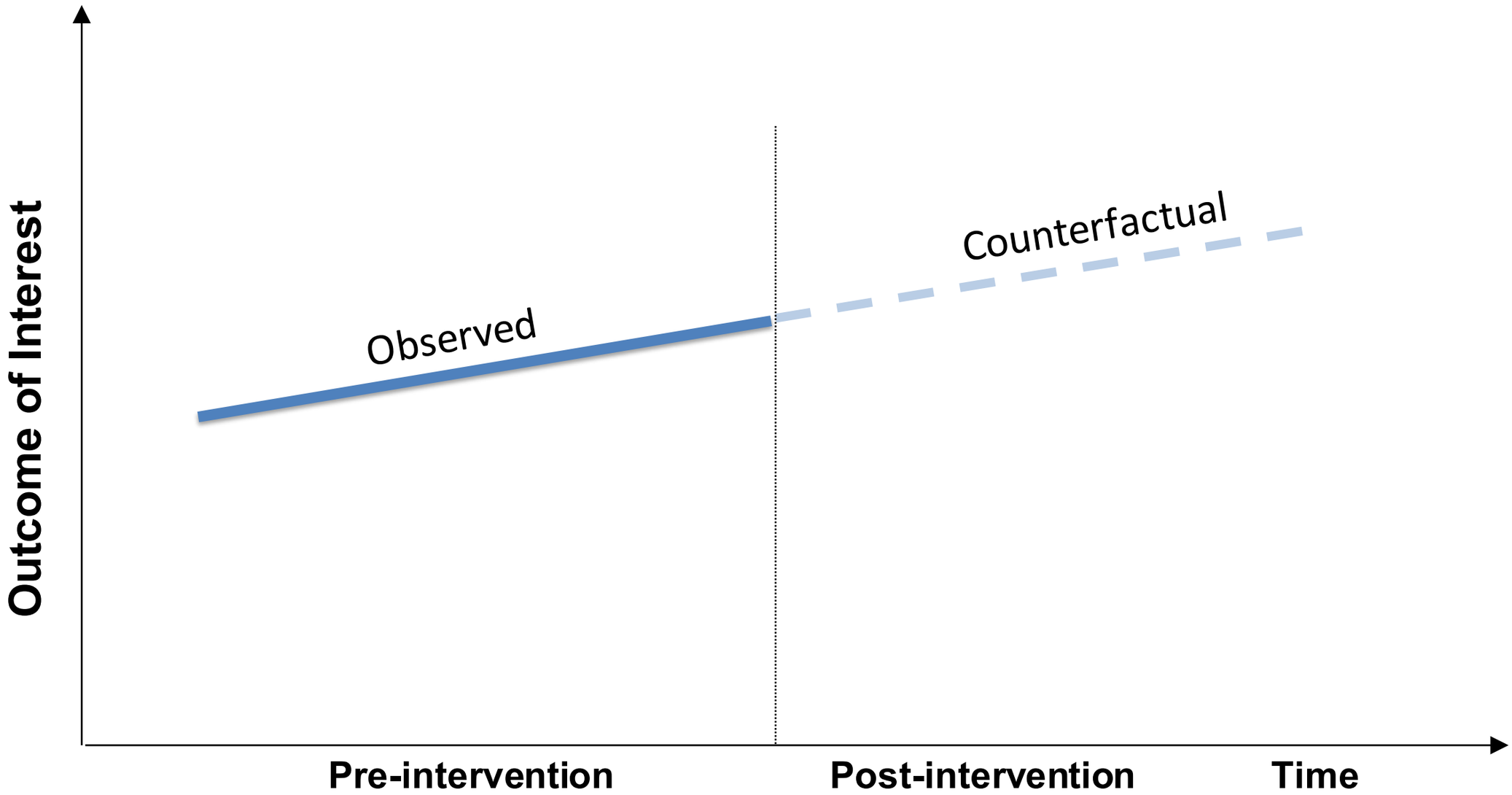

Intervention
Outcome of Interest
Counterfactual
Slope
Change
Observed
Level
Change
Observed
Pre-intervention
Post-intervention
Time


intervention
before
after
No change


intervention
before
after
Level change


intervention
before
after
Trend change


intervention
before
after
Level + trend change


Outcome of Interest
Observed
Counterfactual
Pre-intervention
Post-intervention
Time

Common Time-Series Biases

1. History
   - Something aside from the policy affected the outcome and was implemented near the same time as the policy

2. Instrumentation
   - A change in measurement occurred near the same time as the policy

ITS WITH A CONTROL GROUP

ITS with a Control Group

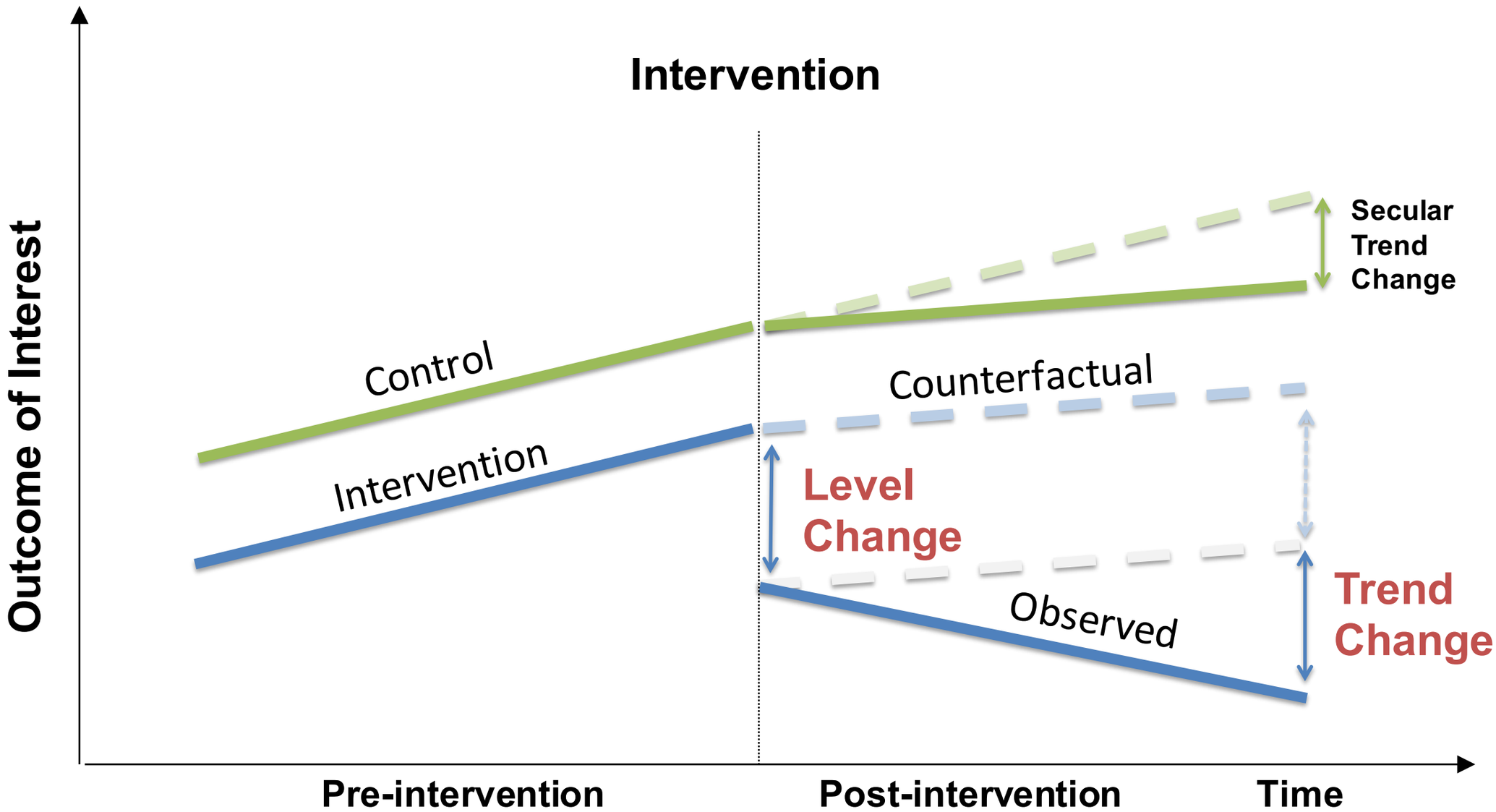
$O_1\ O_2\ O_3\ O_4\ O_5\ O_6\ O_7\ O_8\ X\ O_9\ O_{10}\ O_{11}\ O_{12}\ O_{13}\ O_{14}\ O_{15}\ O_{16}$

$O_1\ O_2\ O_3\ O_4\ O_5\ O_6\ O_7\ O_8\quad O_9\ O_{10}\ O_{11}\ O_{12}\ O_{13}\ O_{14}\ O_{15}\ O_{16}$

Interrupted Time Series with Control

- Design
  - Compare longitudinal trends before and after the policy change between the intervention and control group

- Major Assumption
  - The existing level and trend in the outcome among those exposed to the intervention would have changed identically to the control group absent the intervention


Intervention
Outcome of Interest
Secular
Trend
Change
Control
Counterfactual
Intervention
Level
Change
Observed
Trend
Change
Pre-intervention
Post-intervention
Time

The Control Series

- Counterfactual becomes the observed change in the control group
- Control group adds further legitimacy by limiting possible history threats
- Can be an unaffected group, another jurisdiction, etc.
  - Does not have to have the same pre trend (although a similar trend is more convincing)
  - Also does not have to be balanced

Potential Biases

1. Selection—History
   - Something aside from the intervention affected the outcome and was implemented near the same time as the intervention

2. Selection—Instrumentation
   - A change in measurement for one group occurred near the same time as the intervention

Problems with Interrupted Time Series

- Some potential issues:
  - Requires stable data, over longer time periods
  - Linear trend might not be realistic
  - Requires technical skill to properly fit from a statistical standpoint

- The reward is protection from many threats to validity

Overview of steps

1. Determine time periods
2. Select analytic cohorts
3. Determine outcomes of interest
4. Setup data
5. Visually inspect the data
6. Perform preliminary analysis
7. Check for and address autocorrelation
8. Run the final model
9. Plot the results
10. Predict relative and absolute effects

STEP 1: DETERMINE TIME PERIODS

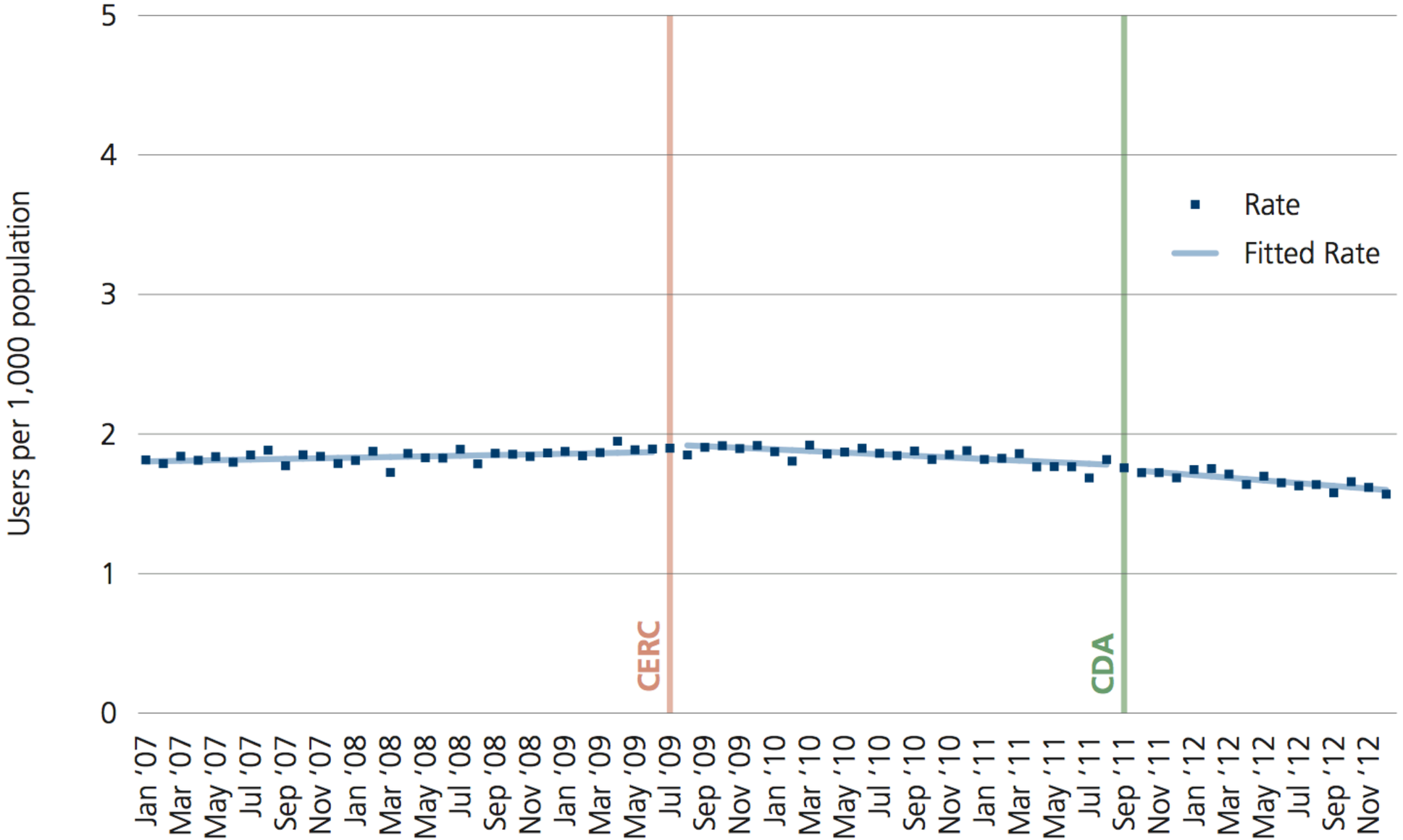
Step 1: Determine time periods

- The “interruption” is the start of the policy or intervention you’re studying

- Can be one point in time, or more than one
  - However, you need enough data between them to measure a trend (8-12 points min)
  - Be wary of confounding interventions

Timing issues to consider

- Length of period
- Anticipatory responses
- Phase-in period
- Co-interventions


Users per 1,000 population
5
4
3
2
1
0
Rate
Fitted Rate
CERC
CDA
Jan '07
Mar '07
May '07
Jul '07
Sep '07
Nov '07
Jan '08
Mar '08
May '08
Jul '08
Sep '08
Nov '08
Jan '09
Mar '09
May '09
Jul '09
Sep '09
Nov '09
Jan '10
Mar '10
May '10
Jul '10
Sep '10
Nov '10
Jan '11
Mar '11
May '11
Jul '11
Sep '11
Nov '11
Jan '12
Mar '12
May '12
Jul '12
Sep '12
Nov '12

STEP 2: SELECT ANALYTIC COHORTS

Step 2: Selecting Analytic Cohorts

- Considerations when selecting groups
  - Where do you expect to see the impact?

- Common to use continuously enrolled, or entire population of an area
  - Problems can arise with attrition from your cohort over time

Control Groups

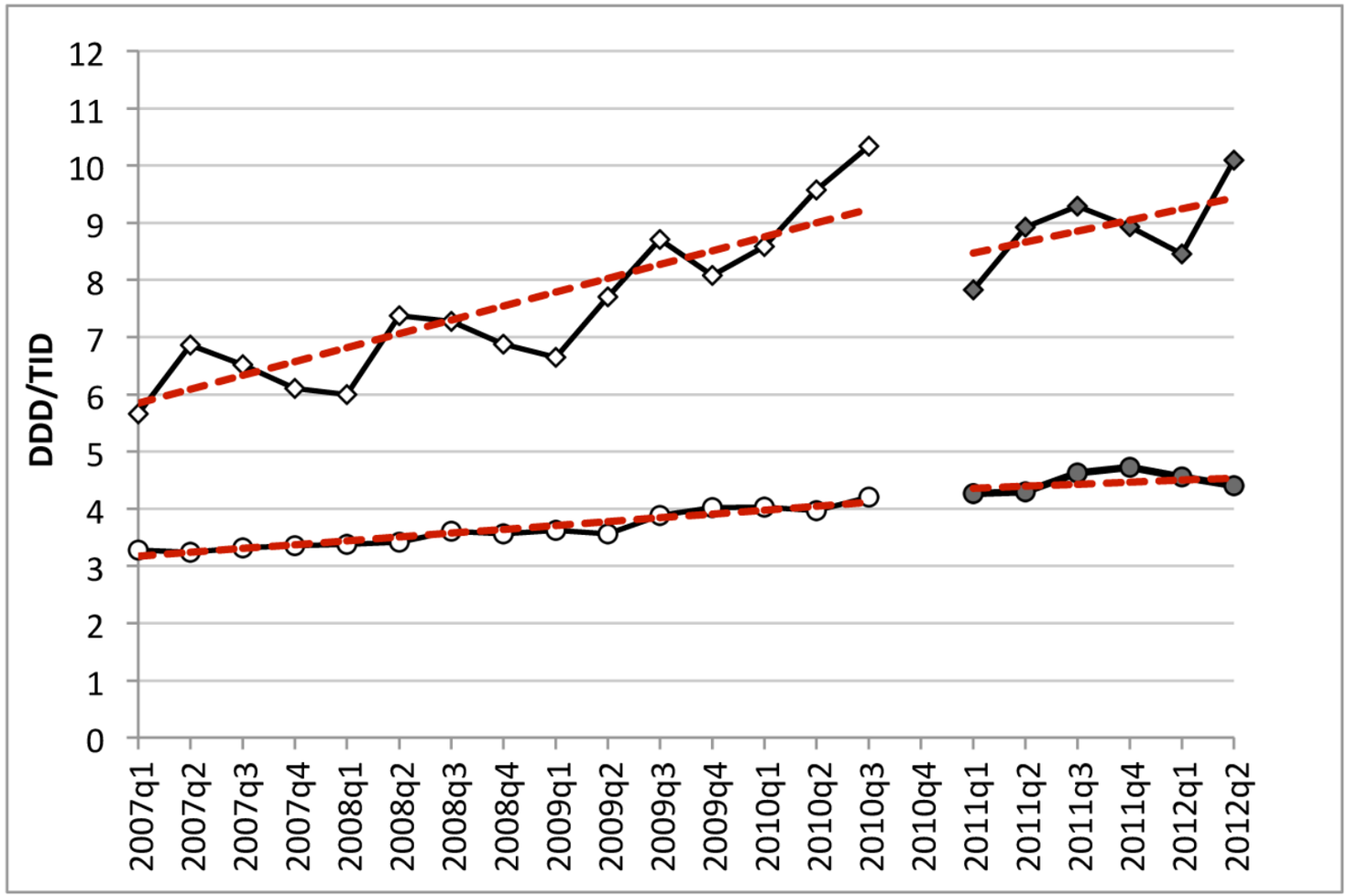
- Control groups can strengthen your inference from an ITS study
  - Could be unexposed group, another region / country, etc.

- Key point: the group can be non-equivalent
  - Doesn't have to "match" the intervention group

Useful Comparison Series

- Similar group not exposed to intervention
  - Another jurisdiction, clinic, group, etc.
  - Can be randomized, matched, or convenience
- Subgroups not expected to be affected

Brazil



DDD/TID antibiotics before the intervention

DDD/TID antibiotics after the intervention

DDD/TID antihypertensives before the intervention

DDD/TID antihypertensives after the intervention

Predicted values

STEP 3: DETERMINE OUTCOMES

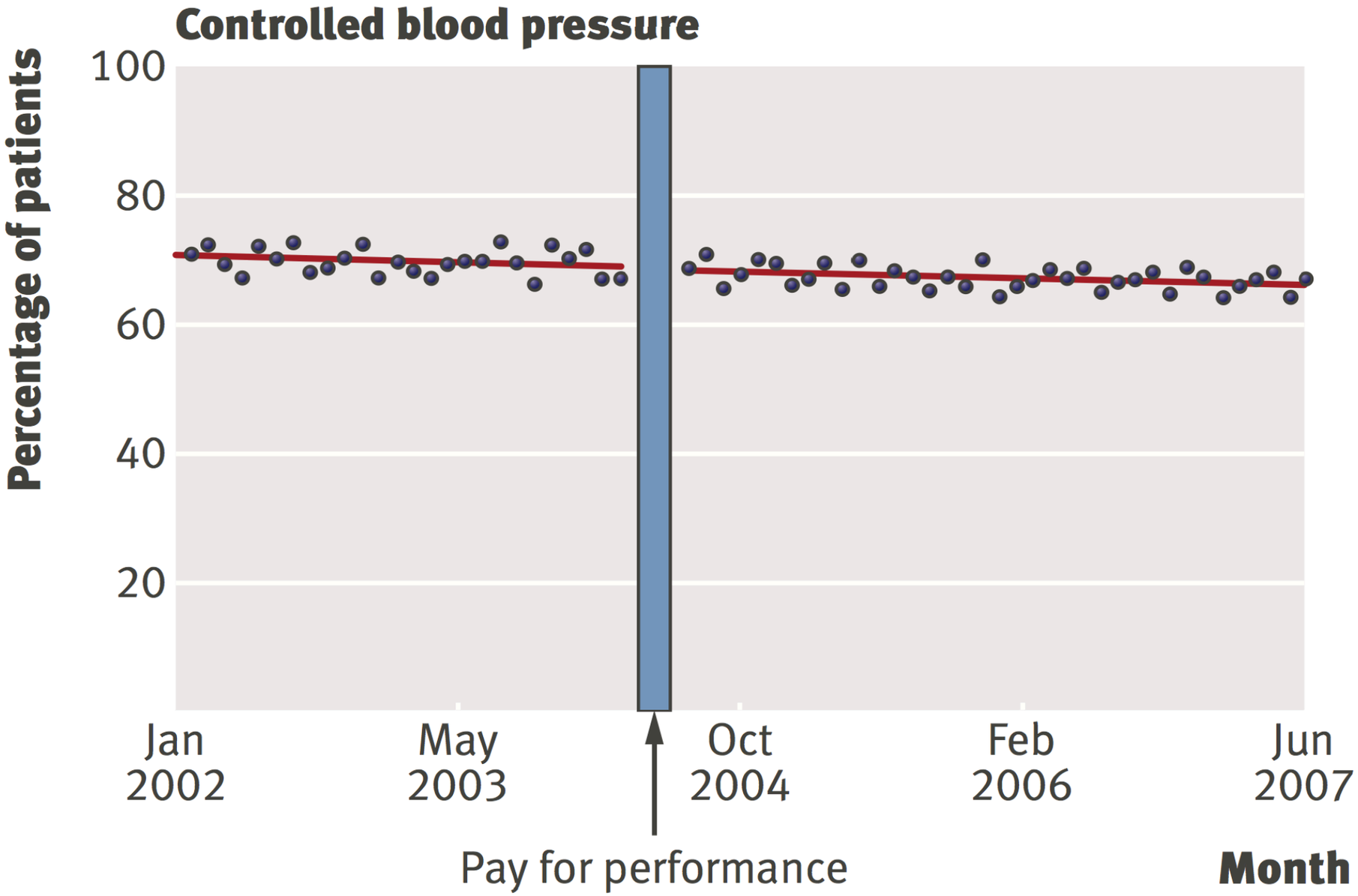
Step 3: Determining Outcomes

- Series of measures of a single characteristic at equidistant time intervals, e.g.:
  - Utilization rates per quarter
  - Costs per month
  - Some other relevant outcome

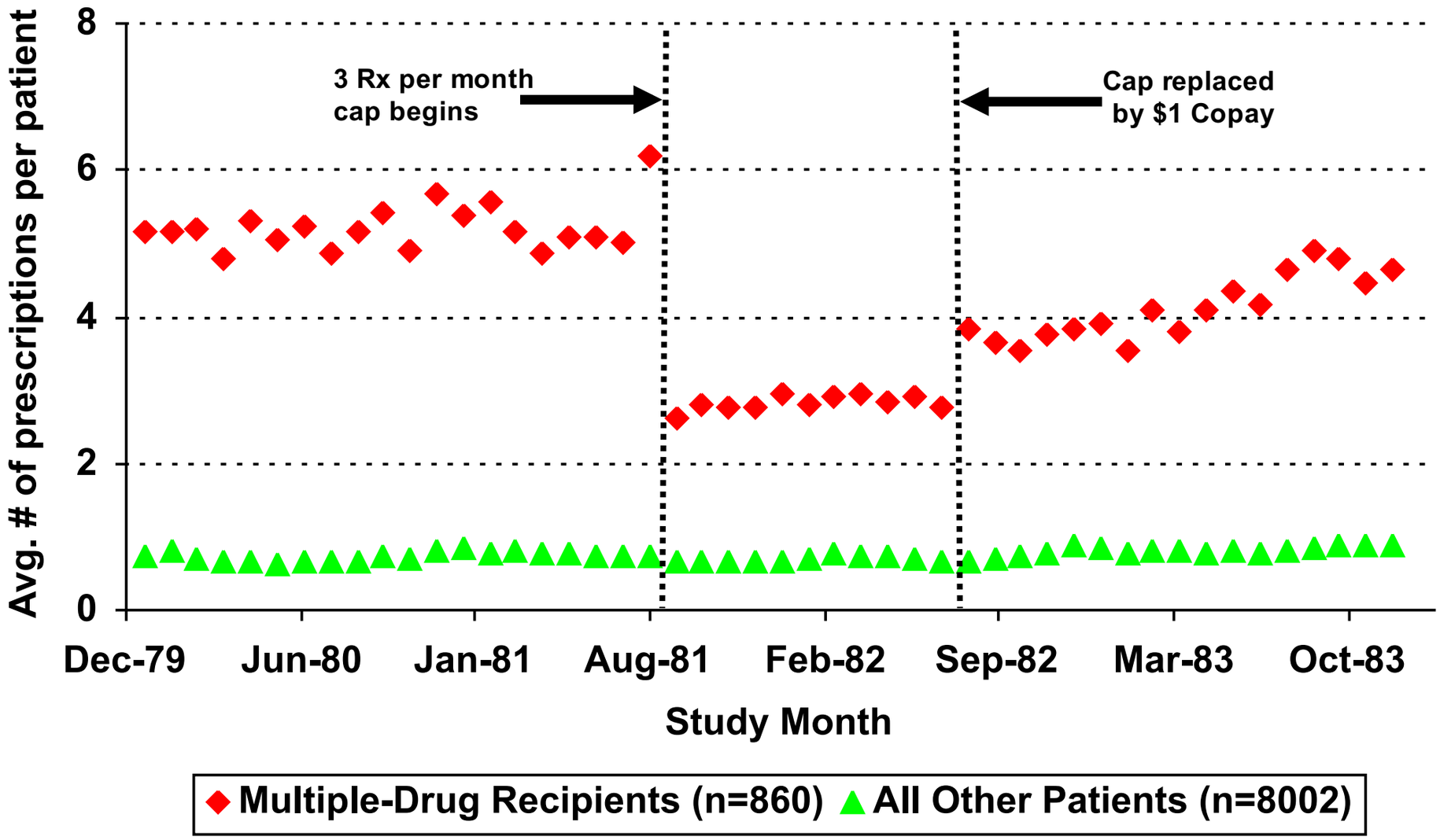
- Reasons these might be collected vary:
  - Evaluating policy changes
  - Routine monitoring or administration

Choosing a Measure

- Choose measures that reflect both the intended and (possibly) unintended outcomes
- Be wary of outcomes that might vary due to attrition
  - If a number of people exit the cohort, this can bias your study
- Rates and proportions often work well



Source: Serumaga B, et al. BMJ 2011; 342:d108



Source: Soumerai et al. NEJM 1987; 317(9):550-6

Overview of steps

1. Determine time periods
2. Select analytic cohorts
3. Determine outcomes of interest
4. Setup data
5. Visually inspect the data
6. Perform preliminary analysis
7. Check for and address autocorrelation
8. Run the final model
9. Plot the results
10. Predict relative and absolute effects