Difference-in-Differences for Social Epidemiology

Danish Epidemiology Society Workshop

Sam Harper

2024-11-06

# *"Methods to Study Social Inequality in Health"*
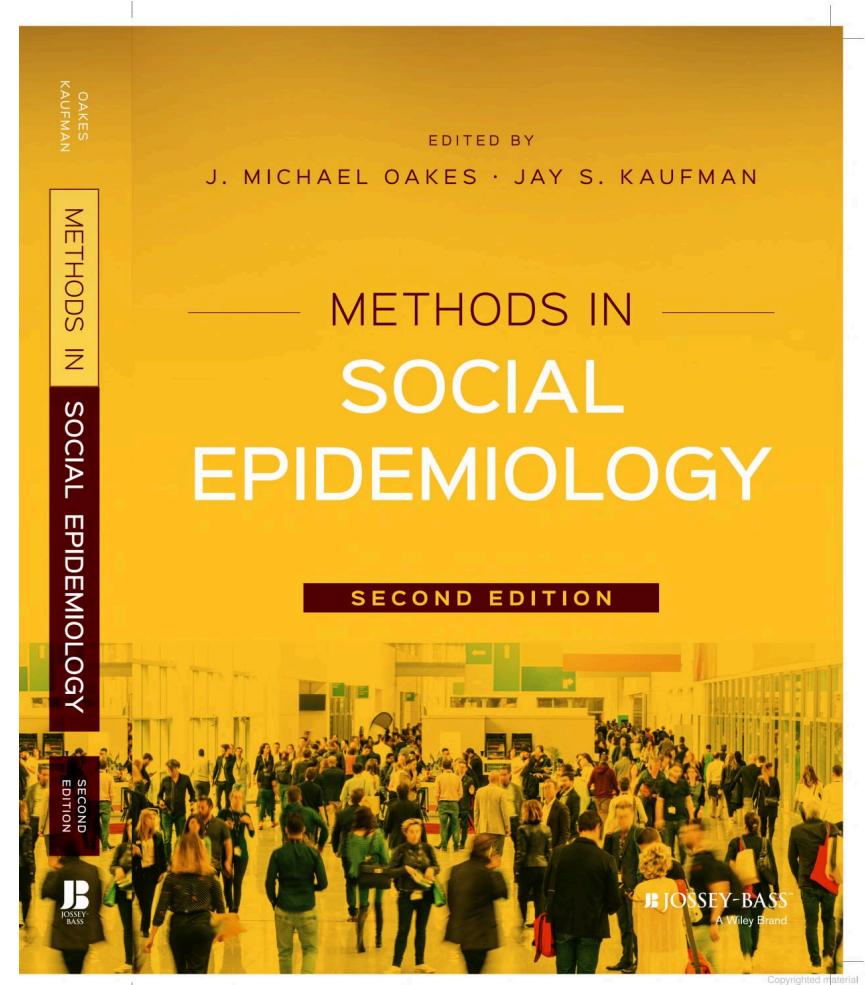
Outline

1. Asking Questions

2. Basics of DD

3. DD for Inequalities

4. Extensions

1. Asking Questions

What are "Methods" in Social Epidemiology?

1. Existence of social differences in health (Descriptive)

2. Causes of observed social differences in health (Etiologic)

   - Overall effect of social group categories (total "effects")

   - Direct/Indirect effects (causal mediation)

3. Policies to address causes and/or remediate social differences in health. (Policy/Intervention)

Where is the evidence?

It is clear that evidence on the effectiveness and cost-effectiveness of public health interventions is often missing. Sometimes this is because policies are insufficiently subjected to outcome evaluation, perhaps because it is assumed that they are mostly beneficial and any positive outcomes can be taken as read.

There is, for example, a wealth of aetiological evidence⋯However, it often appears to be difficult to translate this information into new interventions and even when the interventions are implemented, their evaluation is often problematic.

Petticrew (2007) on "Plugging gaps in the evidence base on health inequalities".

Can interventions make things worse?

> …a large part of the literature is descriptive rather than analytical.

> We found no support for the notion that the methods used to reduce smoking decrease inequalities in health between educational groups.

Vilhelmsson and Östergren (2018)

Reducing health inequalities with interventions targeting behavioral factors among individuals with low levels of education - A rapid review

Andreas Vilhelmsson *[a], **Per-Olof Östergren**[a]

Division of Social Medicine and Global Health, Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden

[a] These authors contributed equally to this work.
* andreas.vilhelmsson@med.lu.se

Abstract

Individuals with low levels of education systematically have worse health than those with medium or high levels of education. Yet there are few examples of attempts to summarize the evidence supporting the efficacy of interventions targeting health-related behavior among individuals with low education levels, and a large part of the literature is descriptive rather than analytical. A rapid review was carried out to examine the impact of such interven-

Causal questions

We want to know:

- Did the program work? If so, for whom? If not, why not?

- If we implement the program elsewhere, should we expect the same result?

- Did it decrease inequalities?

- These questions involve counterfactuals about what would happen *if we intervened to do something*.

- These are causal questions.

Image credit: https://bit.ly/3ApemFr
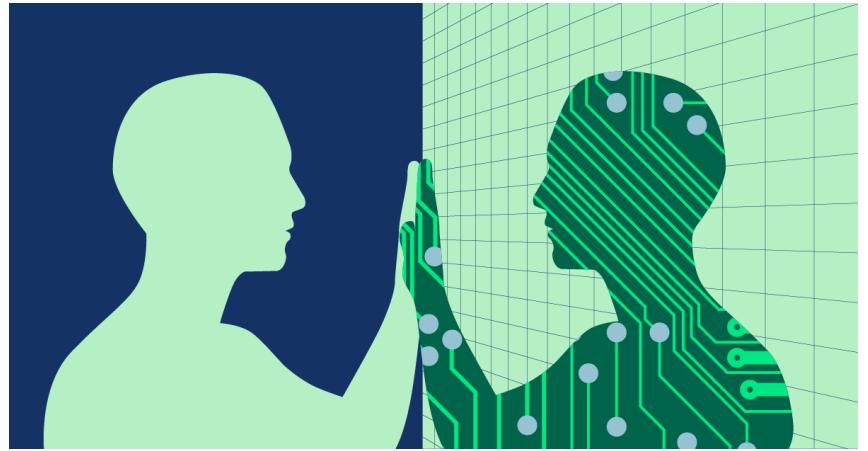
Causation, Association, and Confounding

Causal effect: Do individuals randomly assigned to treatment have better outcomes?

$$E[Y|SET(T=1)] - E[Y|SET(T=0)]$$

Association: Do treated individuals have better outcomes?

$$E[Y|T=1] - E[Y|T=0]$$

Confounding:

$$E[Y|SET(T=1)] - E[Y|SET(T=0)] \neq E[Y|T=1] - E[Y|T=0]$$
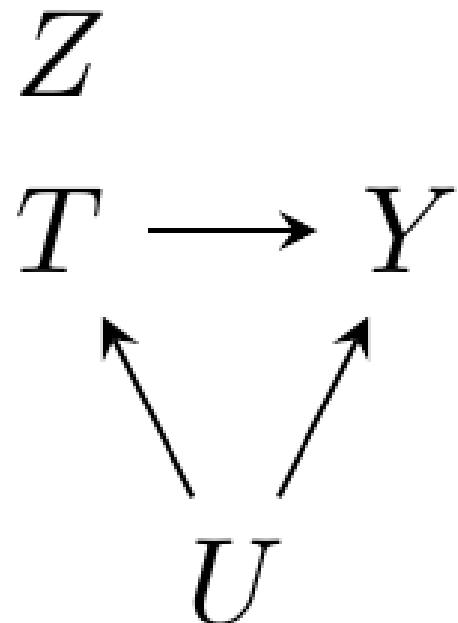
Consequences of non-randomized treatment

- If we aren't controlling treatment assignment, who is?

- Policy programs rarely select people to treat at random.

  - Targeting those most likely to benefit.

  - Programs implemented in response to events.

  - State taxes (or subsidies) for certain goods.

- People do not choose to participate in programs at random.

  - Welfare programs, health screening programs, etc.

  - People who believe they are likely to benefit from the program.

Image credit: Gowtham AGM on Unsplash

Thinking about research design

$$Z$$

$$T \longrightarrow Y$$

$$U$$

Without randomization (Z), we focus on exploiting:
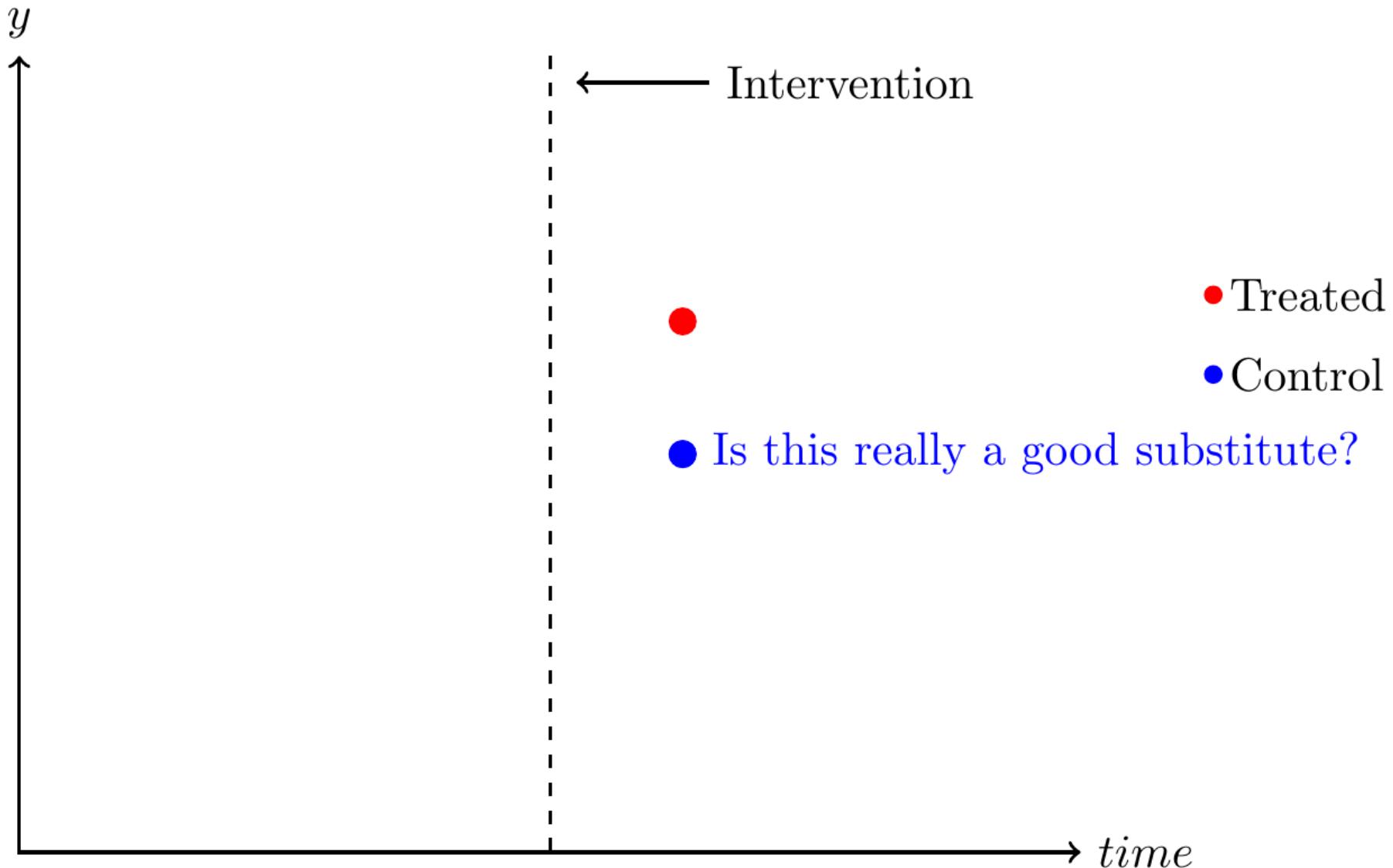
- Treated group with change in exposure.

- Comparison group without change.

Recall the potential outcomes framework. We need a substitute population (treated and controls):
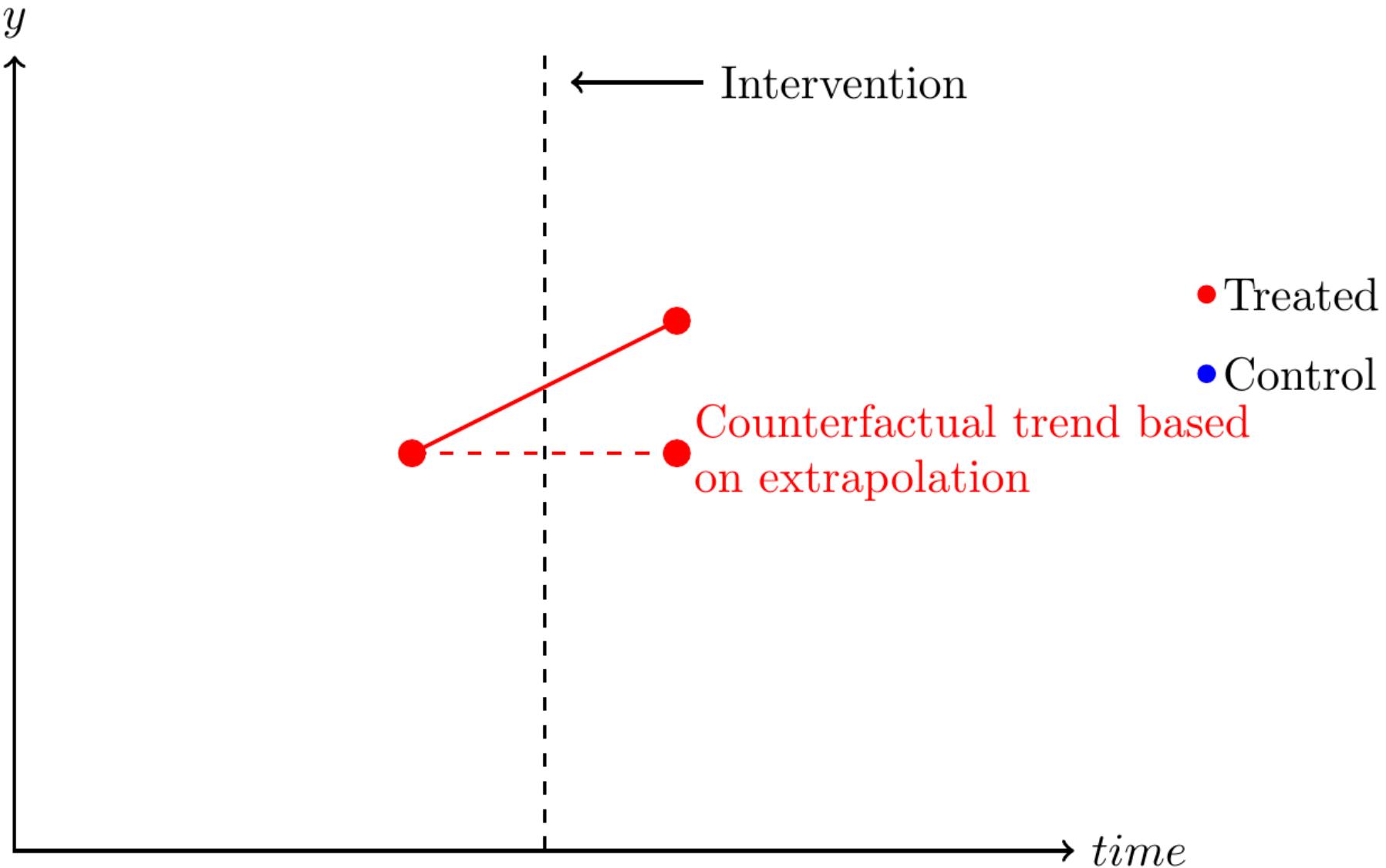
$$E[Y^1 - Y^0] = E[Y^1|T = 1] - E[Y^0|T = 0]$$

- Where should we get our counterfactual?

# One-group posttest design with control group
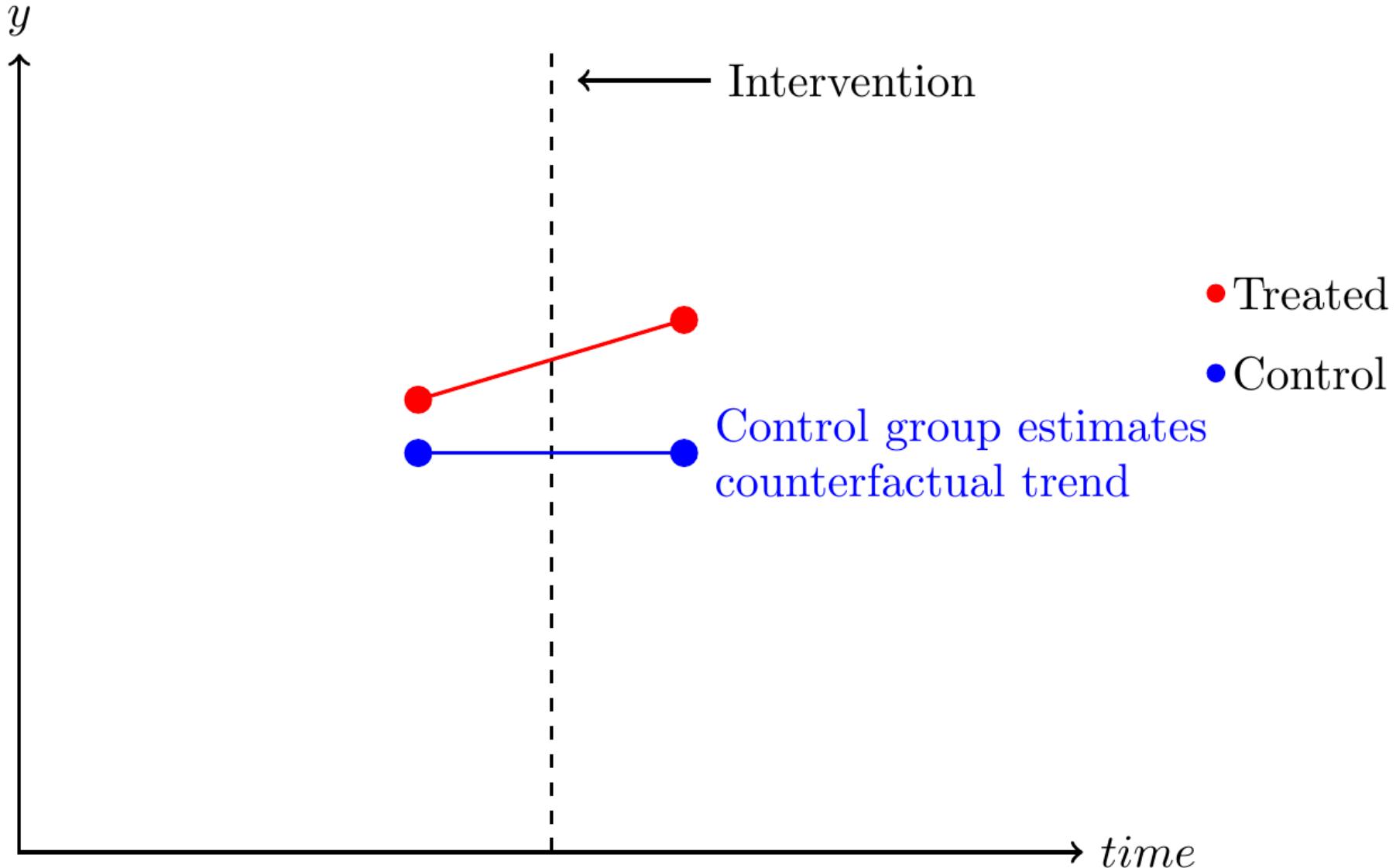
One-group pretest-posttest design

One-group pretest-posttest design

- Even a single pretest observation provides some improvement over the posttest only design.

- Now we derive a counterfactual prediction from the same group before the intervention.

- Provides weak counterfactual evidence about what would have happened in the absence of the program.

  - We know that Y_{t-1} occurs before Y_t (correct temporal ordering).

  - Could be many other reasons apart from the intervention that $Y_t \neq Y_{t-1}$.

- Stronger evidence if the outcomes can be reliably predicted and the pre-post interval is short.

- Better still to add a pretest and posttest from a control group.
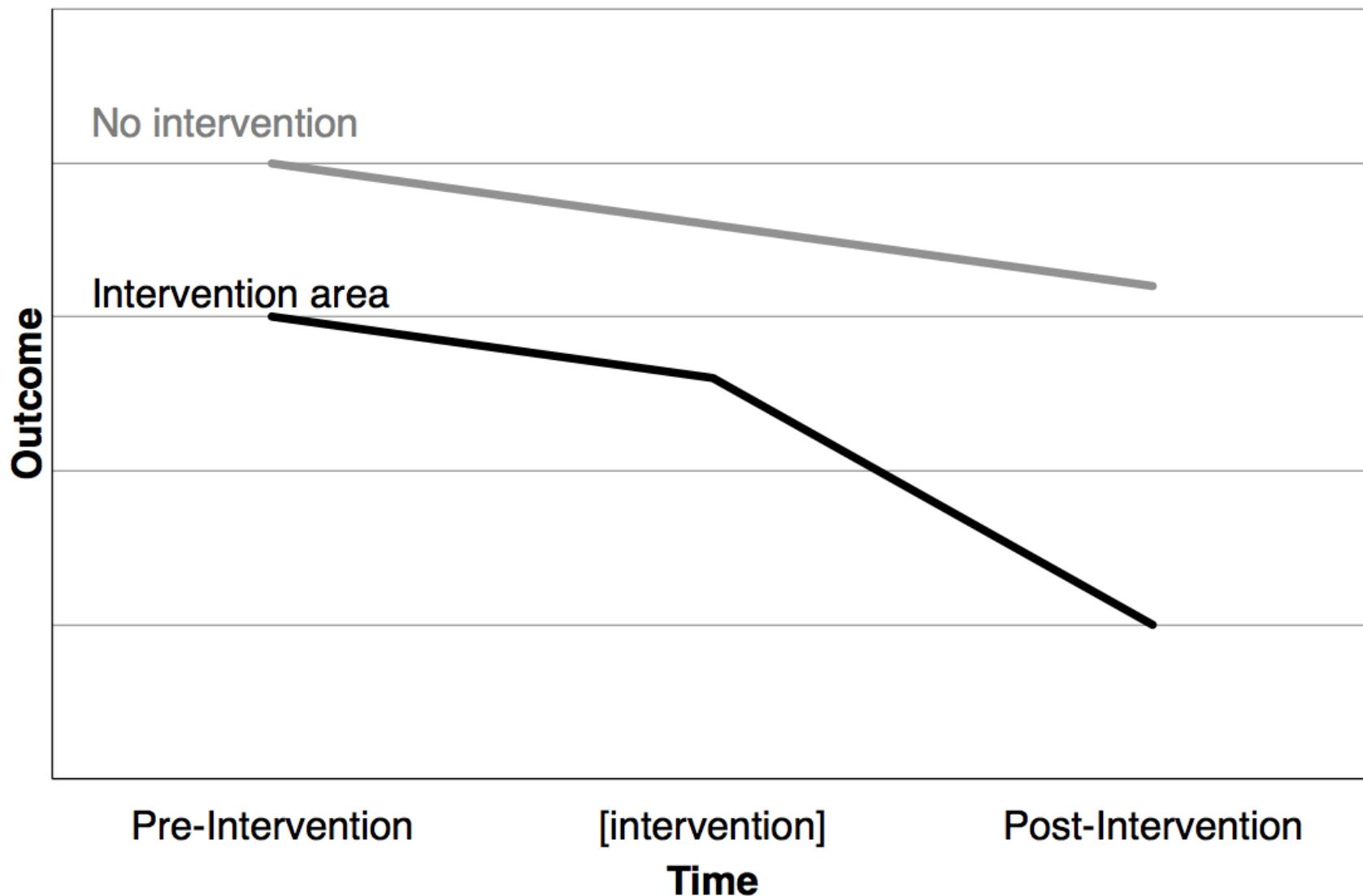
Adding pretests for both groups
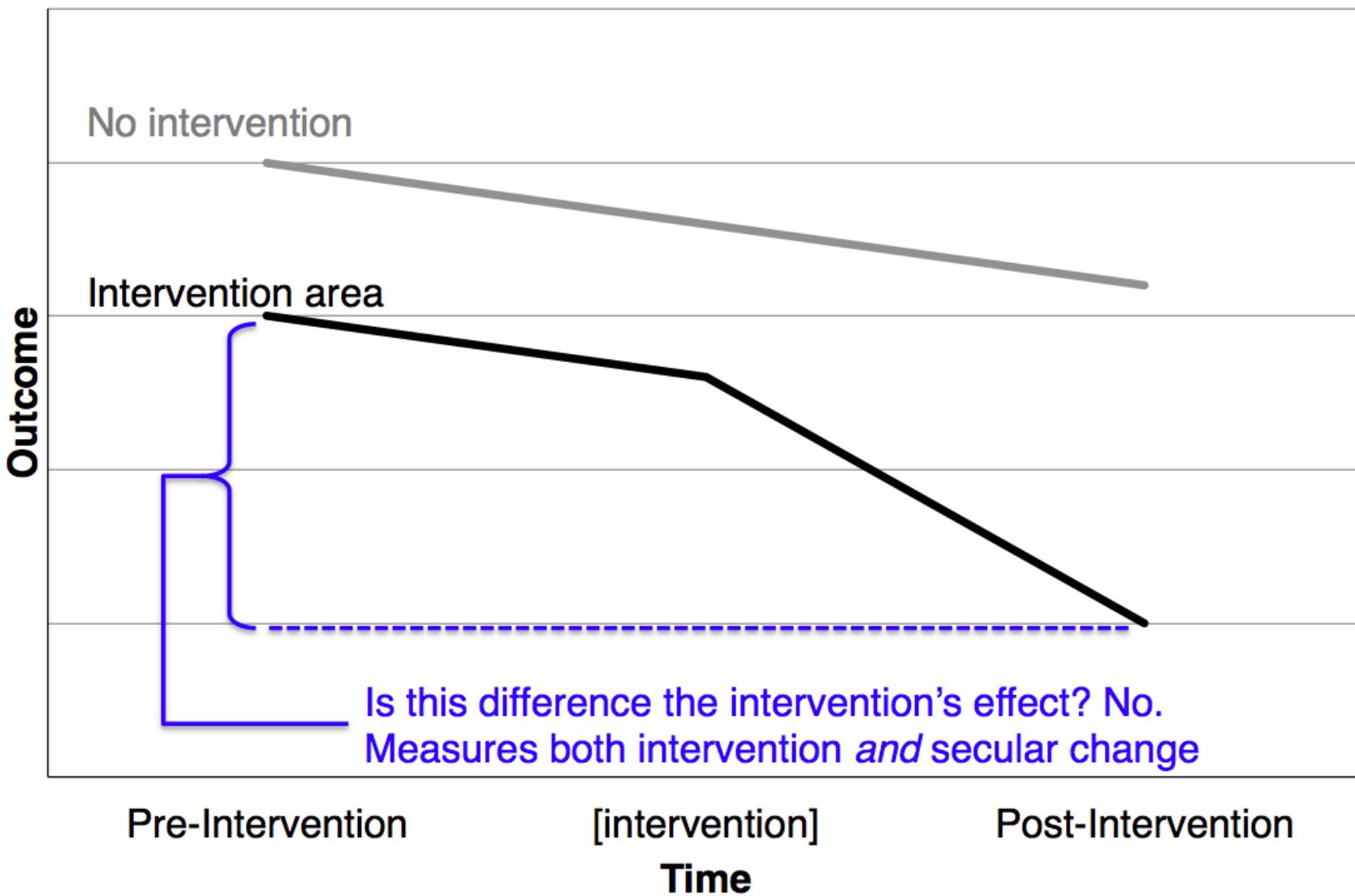
How does this help?

- Pre/post in control helps resolve this by differencing out any time-invariant characteristics of both groups.

  - Many observed factors don't change over the course of an intervention (e.g., geography, parents' social class, birth cohort).

  - Any time-invariant *unobserved* factors also won't change over intervention period.

  - We can therefore effectively control for them.

- Measuring same units before and after a program cancels out any effect of all of the characteristics that are unique to that observation and that do not change over time.

- This also has the benefit of canceling out (or controlling for) unobserved time-invariant characteristics.
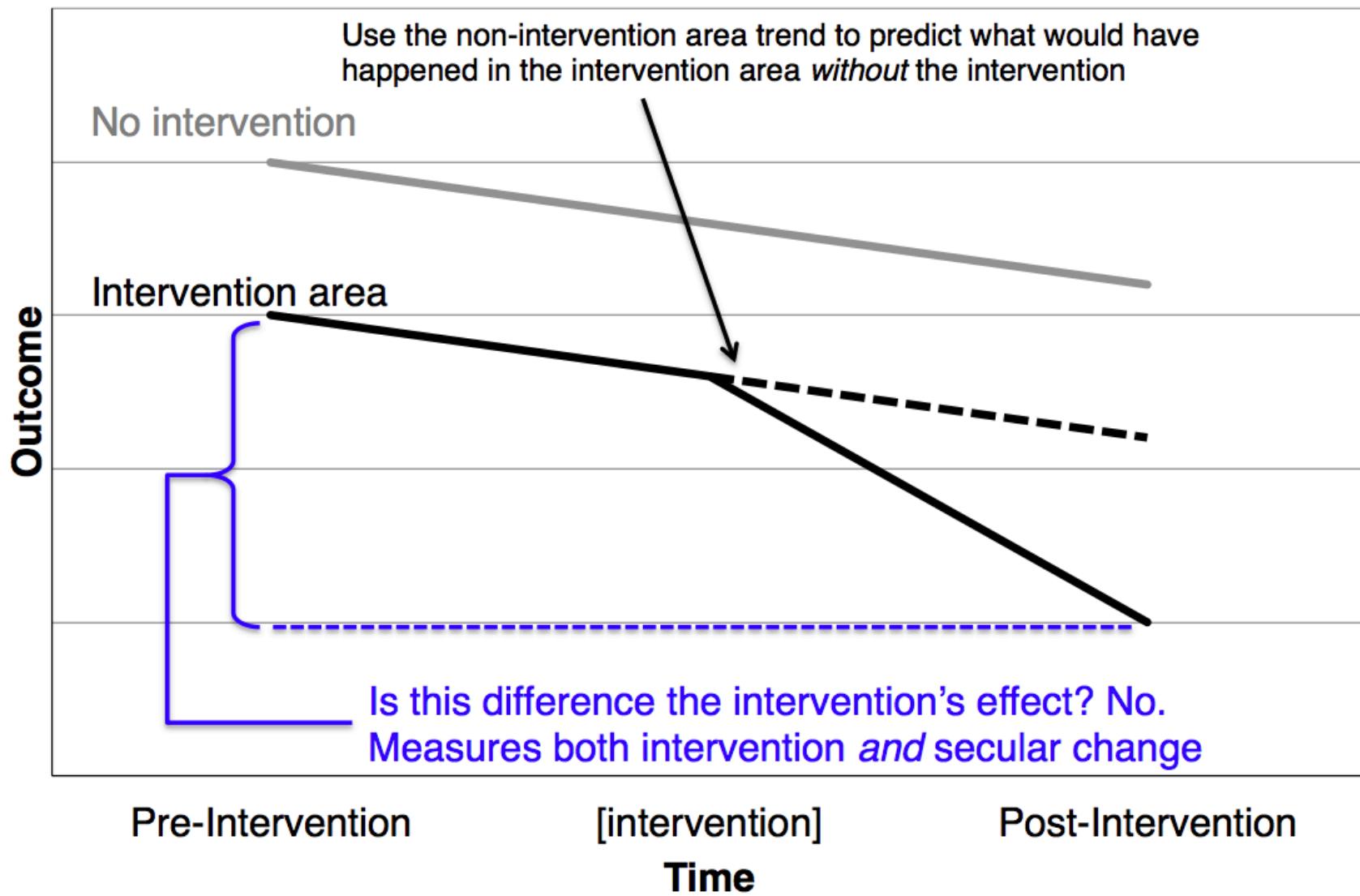
The need for a control group
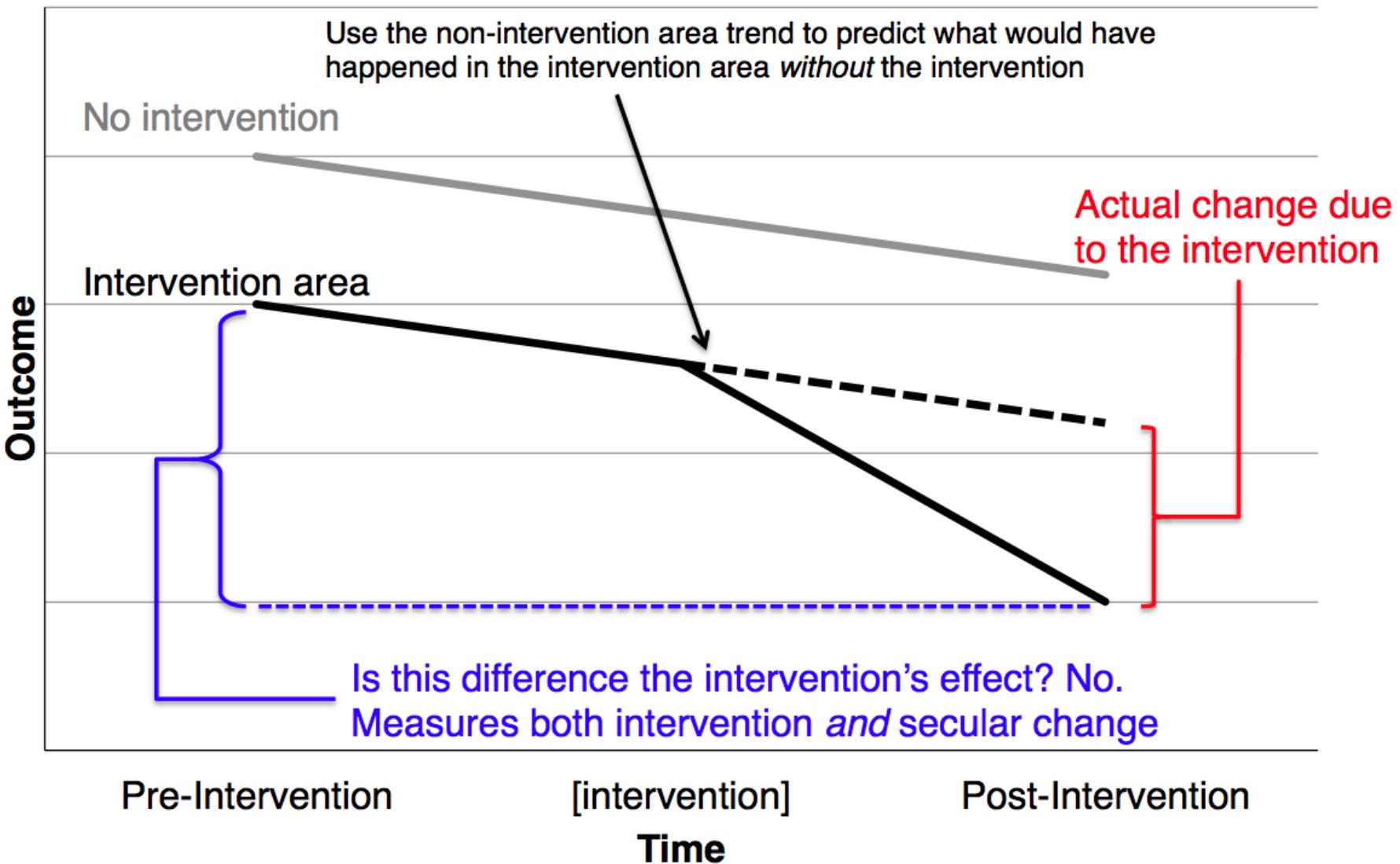
What if there were no intervention?
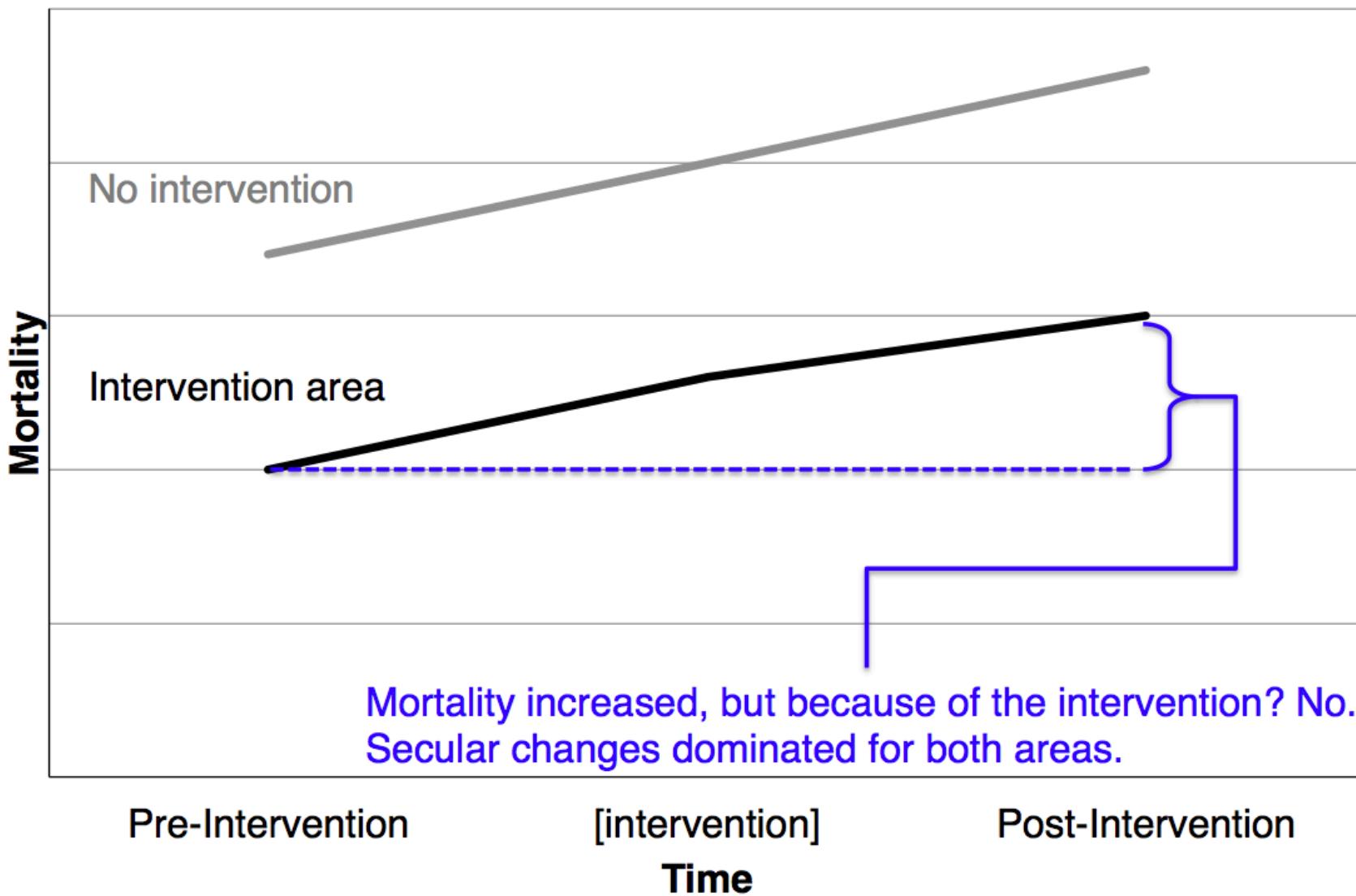
What if there were no intervention?
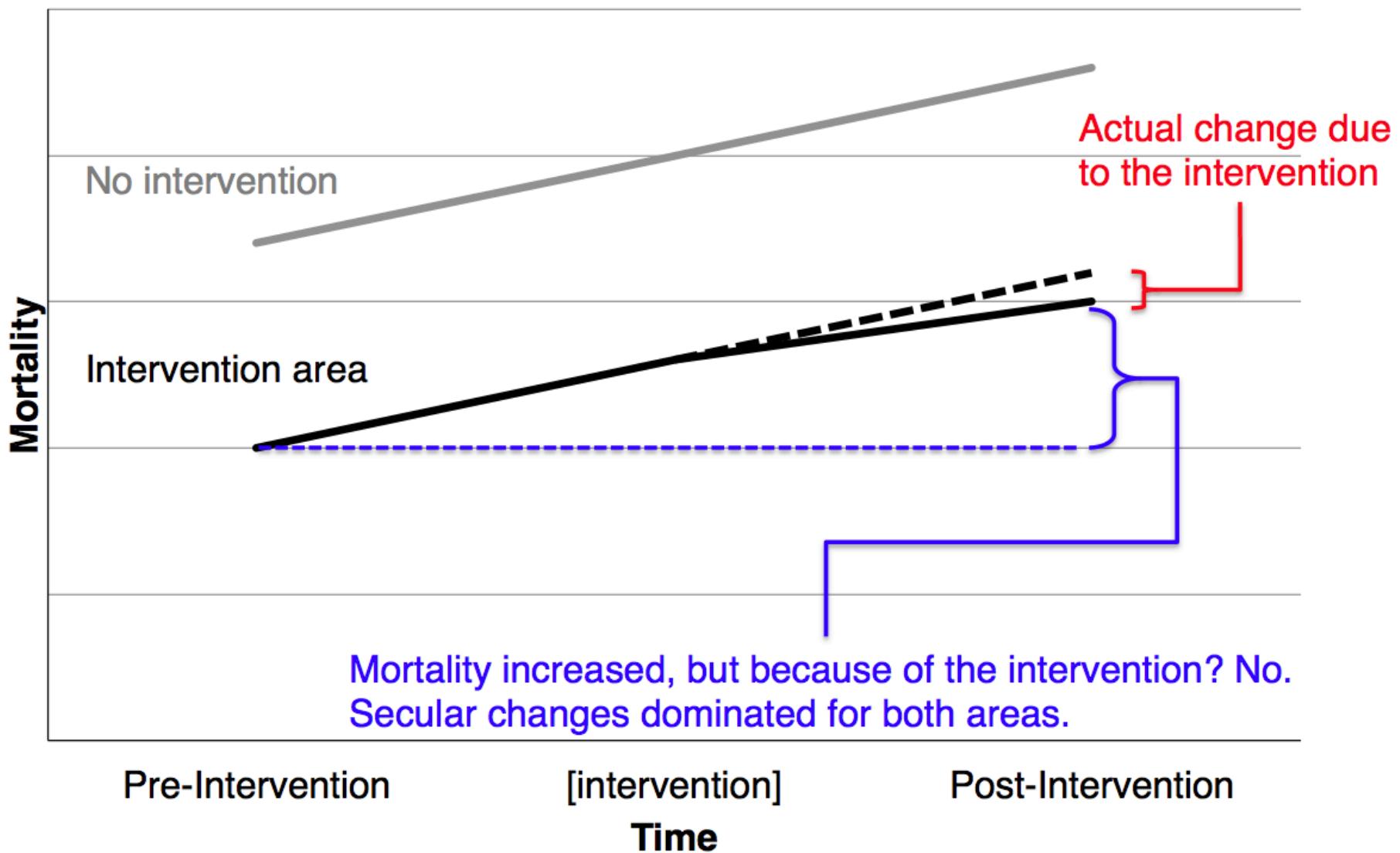
What if there were no intervention?
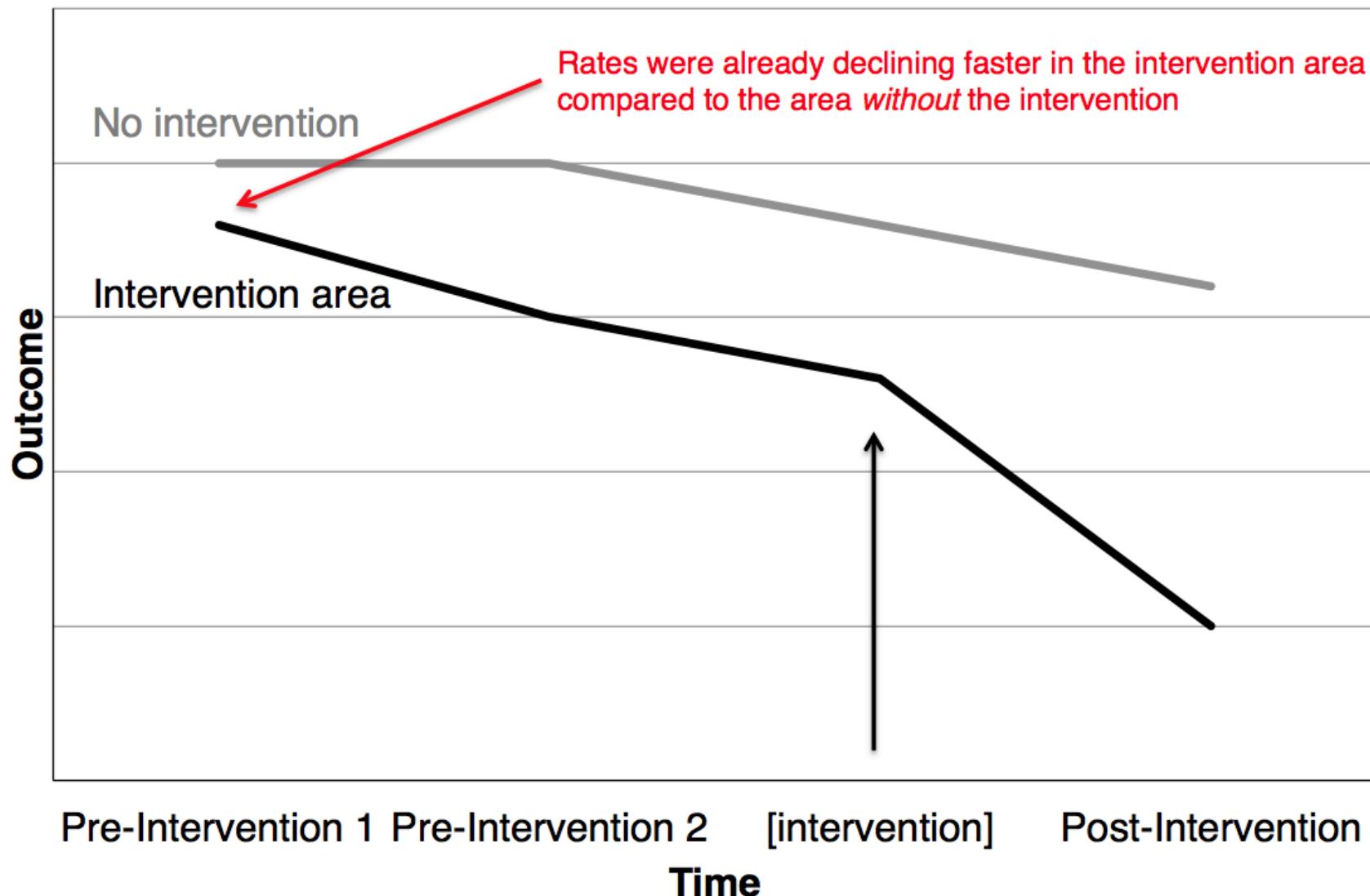
What if things worsened over time?
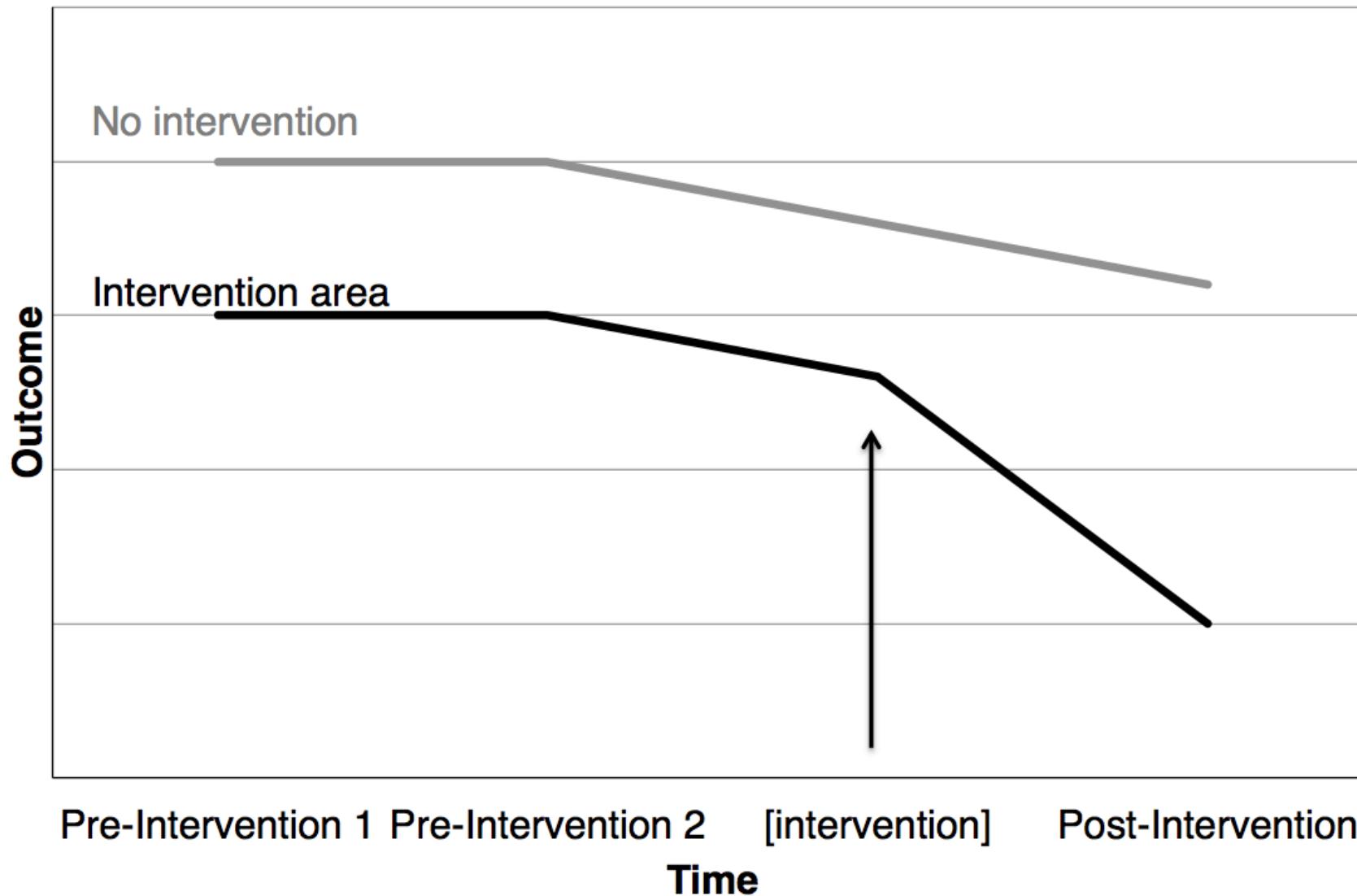
What if things worsened over time?

More time periods make better evaluations

Rates were already declining faster in the intervention area compared to the area *without* the intervention

No intervention

Intervention area

Outcome

Pre-Intervention 1 Pre-Intervention 2 [intervention] Post-Intervention
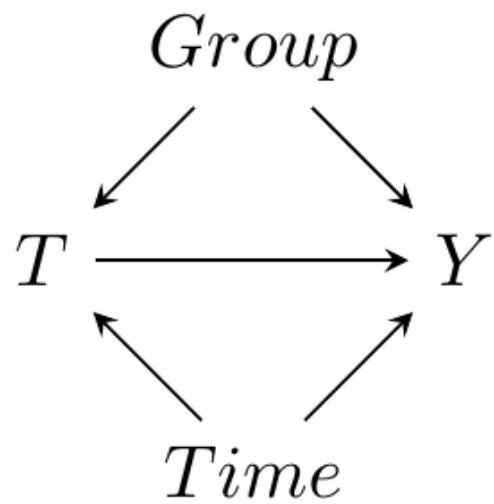
Time

Parallel pre-trends increase "exchangeability"

2. Basics of DD
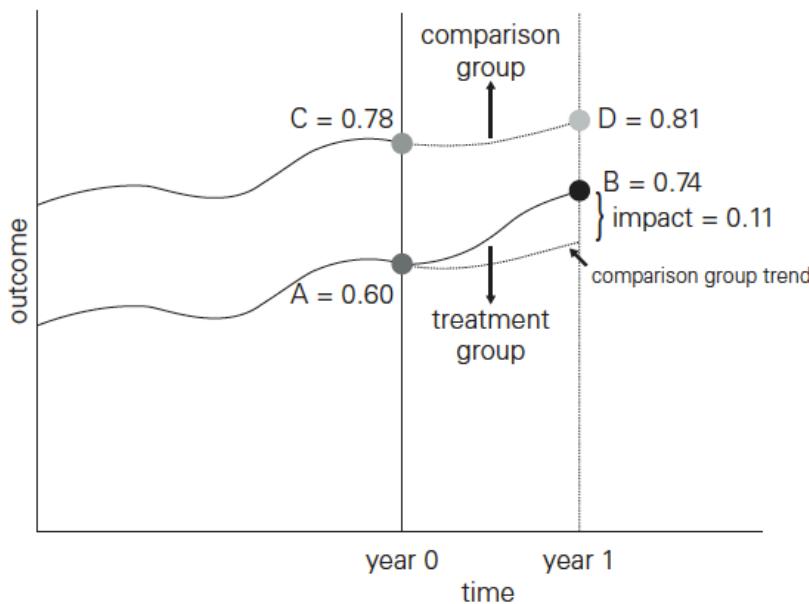
Difference-in-Differences: Basic Idea

The simplest DD setting:
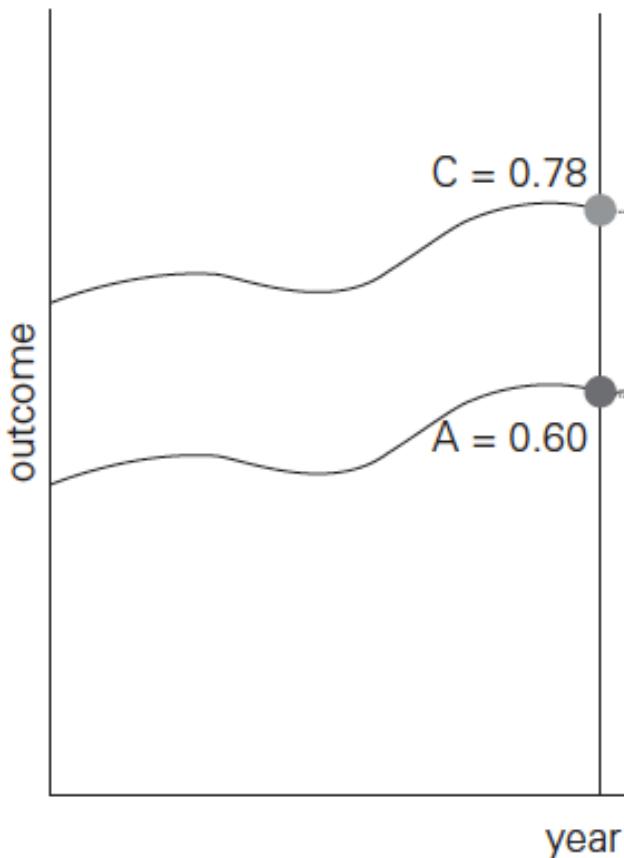
- Outcomes observed for "units" observed in one of two groups:

  - Treated

  - Control

- Outcomes observed in one of two time periods:

  - Before intervention

  - After intervention.

- Treated: only units in one of the two groups are exposed to a treatment, in the second time period.

- Control: Never observed to be exposed to the treatment.

Difference-in-Differences: Basic Idea

- The average change over time in the non-exposed (control) group is subtracted from the change over time in the exposed (treatment) group.

- Double differencing removes biases in second period comparisons between the treatment and control group that could result from:

  - Fixed (i.e., non time-varying) differences between those groups.

  - Comparisons over time in the treatment group that could be the result of time trends unrelated to the treatment.
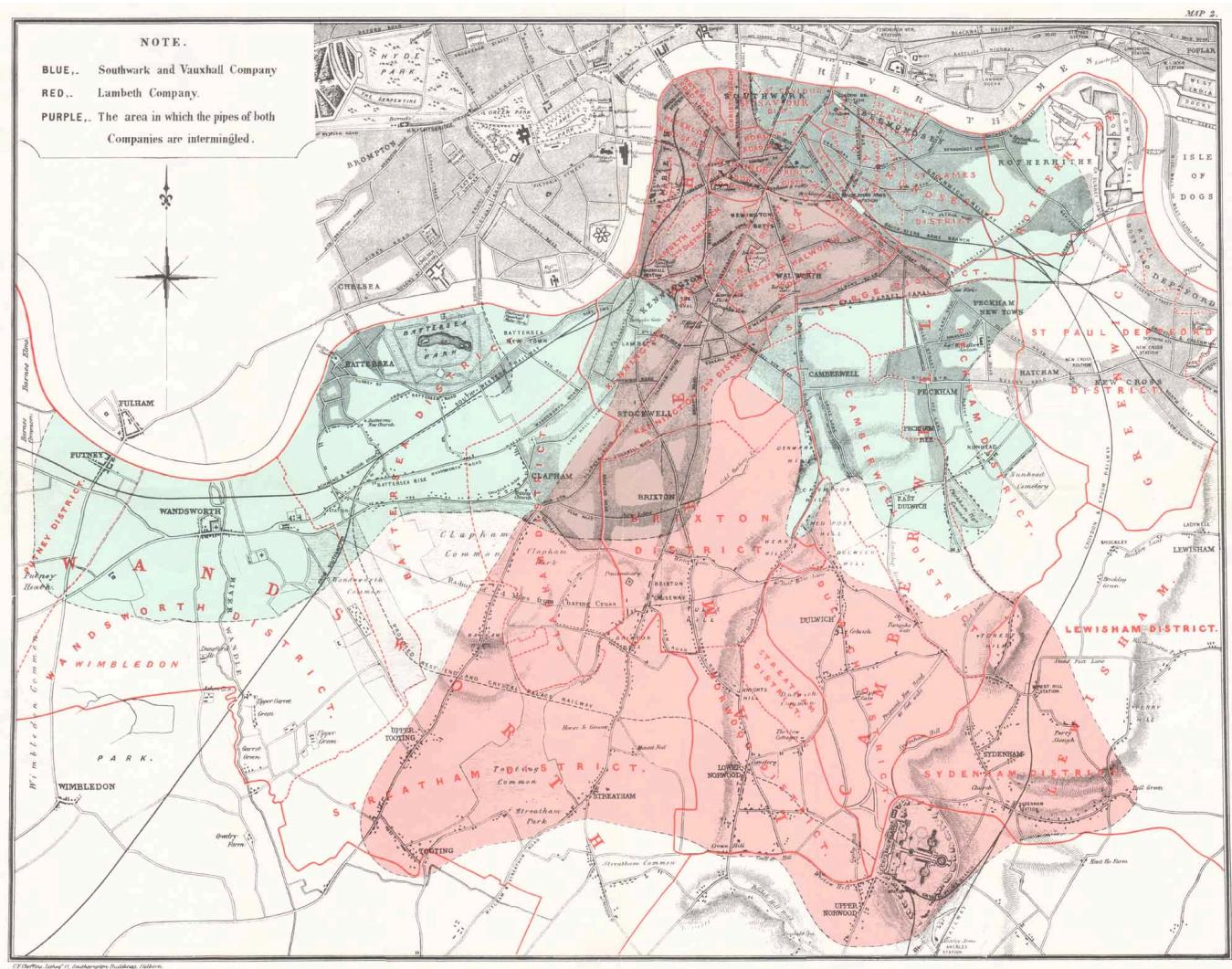
Key Assumption: Parallel Trends

- Basic DD controls for any time invariant characteristics of both treated and control groups.

- Does not control for any time-varying characteristics.

- If another policy/intervention occurs in the treated (or control) group at the same time as the intervention, we cannot cleanly identify the effect of the program.

- DD main assumption: in the absence of the intervention treated and control groups would have displayed similar trends.

- This is called the *parallel trends* assumption.

Impossible to verify, see Gertler et al. (2016).

Classic epidemiology example: Water and cholera
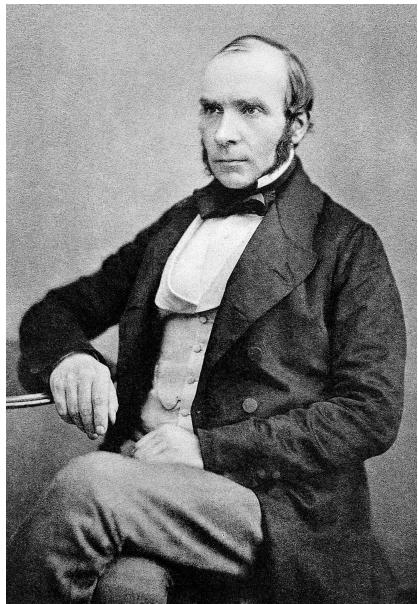
See Snow (1855) reprinted as Snow et al. (1936)

Snow's method

- Couldn't randomize.
- Lambeth moved intake upstream of London *after* 1849.
- SV similar to Lambeth, but did not move.
- SV as 'unaffected' control.
- Did not estimate DD parameter, but idea was there.

Region	Rate (1849)	Rate (1854)	Post-Pre
Lambeth (treated)	130.1	84.9	-45.2
Southwark + Vauxhall (control)	134.9	146.6	11.7
Group Diff (treat - control)	-4.8	-61.7	???

Why is Snow's work compelling?

- Evidence of pre-treatment equivalence between groups:

"In many cases a single house has a supply different from that on either side. Each company supplies both rich and poor, both large houses and small; there is no difference either in the condition or occupation of the persons receiving the water of the different companies···"

- No knowledge of mechanisms or intervention:

"divided into two groups without their choice, and, in most cases, without their knowledge"

See Snow (1855) reprinted as Snow et al. (1936) and also Freedman (1991)

Difference-in-Differences without Regression

DD (can be) just differences in means.

Let $\mu_{it} = E(Y_{it})$

- $i = 0$ is control, $i = 1$ treated.
- $t = 0$ is pre, $t = 1$ is post.
- One 'difference' is pre-post in treated:
  $\mu_{11} - \mu_{10}$
- Second 'difference' is pre-post in control:
  $\mu_{01} - \mu_{00}$
- Differences-in-Differences:
  $(\mu_{11} - \mu_{10}) - (\mu_{01} - \mu_{00})$

Snow's Example:

Area	Pre	Post	Difference
Treated	130	85	-45
Control	135	147	12
T - C	-5	-62	-57

DD Regression: Two Groups, Two Periods (2x2)
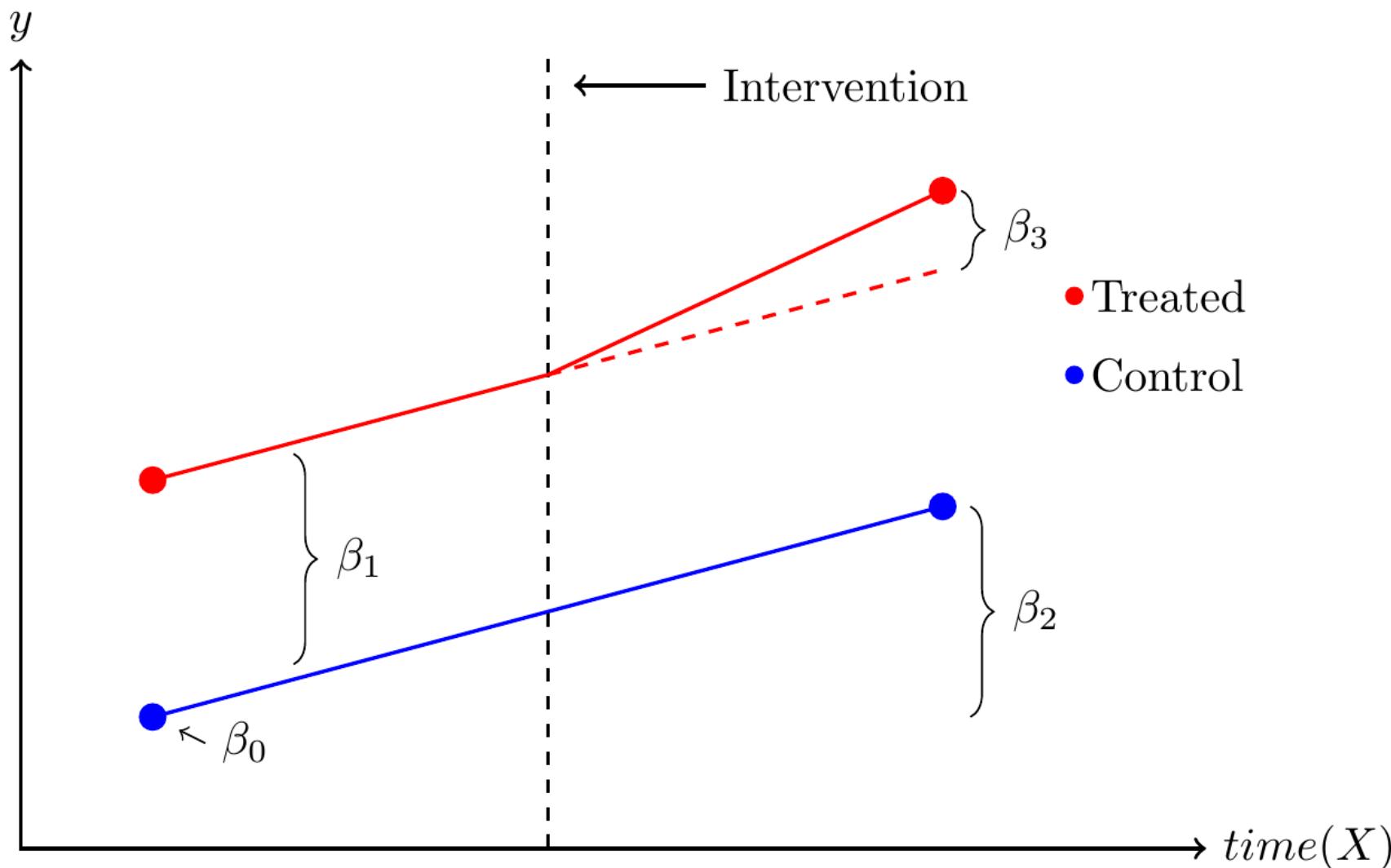
Single treated and control group, two periods:

- β_1 = Treated group
- β_2 = Post period
- β_3 = Product term

y	group	time	treat?	post?	treatXpost
1	1	0	0	0	
1	2	0	1	0	
2	1	1	0	0	
2	2	1	1	1	

$$Y = \beta_0 + \beta_1 * treat + \beta_2 * post + \beta_3 * treat * post$$

$$Y = \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 Treat * Post + \varepsilon_t$$

Difference-in-differences estimates the ATT

Focus on treated group due to selection

- Our DD model is: $Y = \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 Treat * Post + \varepsilon$.

- In the (possibly counterfactual) absence of intervention, the expected outcome is:
  - $E\left(Y_i^0 | T = 1, A = 0\right) = \beta_1 + \beta_2$

- In the (possibly counterfactual) presence of intervention, the expected outcome is:
  - $E\left(Y_i^1 | T = 1, A = 1\right) = \beta_1 + \beta_2 + \beta_3$

- ATT is the expected difference in $Y_i^1 - Y_i^0$ for those treated in the post-period:
  - $ATT = E\left(Y^1 - Y^0 | T = 1\right) = \beta_3$

Back to Snow's water pump

How to estimate the impact?

- $Treat = 1$ if Lambeth, 0 if SW
- $Post = 1$ if 1854, 0 if 1849
- $Treat * Post = 1$ if Lambeth in 1854, 0 otherwise.

$$Y = \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 Treat * Post + \varepsilon$$

Province, Time	Estimate	Time Diff	DD
SW, 1849	β_0		
		β_2	
SW, 1854	$\beta_0 + \beta_2$		
			β_3
Lambeth, 1849	$\beta_0 + \beta_1$		
		$\beta_2 + \beta_3$	
Lambeth, 1854	$\beta_0 + \beta_1 + \beta_2 + \beta_3$		

Reformulation of the model using 'fixed effects'

Express our earlier model using 'fixed effects':

- Dummy for Group
- Dummy for Time
- *Time-varying* policy indicator

$$Y = \beta_0 + \beta_1 * Group2 + \beta_2 * Time2 + \beta_3 * policy$$

y	group	time	treat?	post?	treatXpost	Group 2	Time 2	policy
1	1	1	0	0	0	0	0	0
1	2	0	1	0	0	1	0
2	1	1	0	0	1	0	0
2	2	1	1	1	1	1	1

β_3 still estimates the 'difference-in-differences' parameter.

What about multiple treated groups?

- Easy to rewrite our earlier model for multiple groups treated at the same time.
- 3 units and 3 time periods.
- Groups 1 and 3 implement policy at T2.
- g2 and g3 are dummies for group 2 and 3
- t2 and t3 are respective time dummies for periods 2 and 3.

y	group	time	policy	g2	g3	t2	t3
.	1	1	0	0	0	0	0
.	1	2	1	0	0	1	0
.	1	3	1	0	0	0	1
.	2	1	0	1	0	0	0
.	2	2	0	1	0	1	0
.	2	3	0	1	0	0	1
.	3	1	0	0	1	0	0
.	3	2	1	0	1	1	0
.	3	3	1	0	1	0	1

Extending the model to multiple groups/times

- The regression model with group and time fixed effects would now look something like this (where β_5 is the DD estimate where policy=1 vs. 0):

$$Y_{gt} = \beta_0 + \beta_1 g2 + \beta_2 g3 + \beta_3 t2 + \beta_4 t3 + \beta_5 p_{gt} + \varepsilon_{st}$$

- Reference categories (for interpreting β_0) are group 1 $(g1)$ and time 1 $(t1)$.

- More generally, you could write the basic equation with multiple group (γ_g) and time (τ_t) fixed effects as:

$$Y_{gt} = \alpha + \gamma_g + \tau_t + \delta^{DD} p_{gt} + \varepsilon_{st}$$

where δ^{DD} is the difference-in-differences estimate for groups treated at time t.

3. DD for Inequalities

Evaluating impact on inequalities
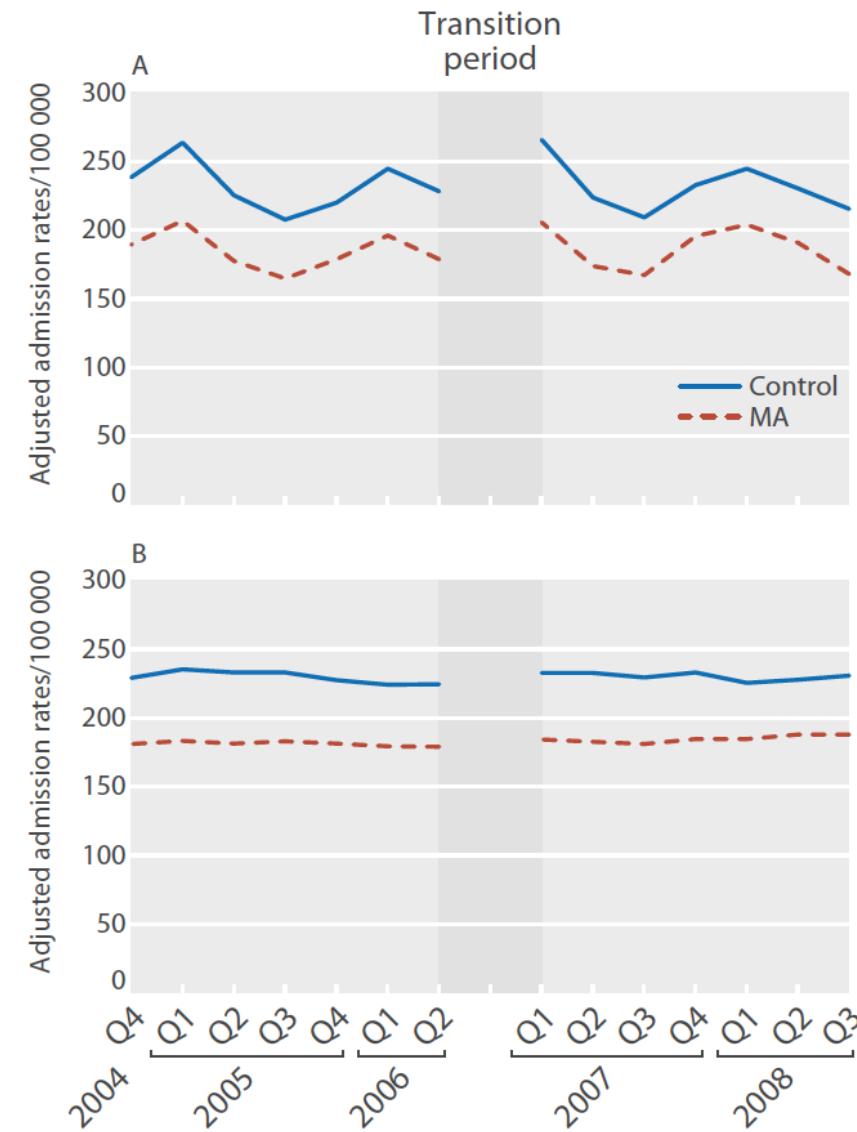
Effect of Massachusetts healthcare reform on racial and ethnic disparities in admissions to hospital for ambulatory care sensitive conditions: retrospective analysis of hospital episode statistics

Danny McCormick,[1] Amresh D Hanchate,[2,3] Karen E Lasser,[3] Meredith G Manze,[3] Mengyun Lin,[3] Chieh Chu,[3] Nancy R Kressin[2,3]

- Evaluated impact of MA reform on inequalities in hospital admissions.

- Compared MA to nearby states: NY, NJ, PA.

- Intervention "worked": % uninsured halved (12% to 6%) from 2004-06 to 2008-09.

McCormick et al. (2015)

We want credible counterfactuals

- Strong visual evidence that pre-intervention trends similar in treated and control groups.

- Adds credibility to assumption that post-intervention trends would have been similar in the absence of the intervention.

McCormick et al. (2015)

Little evidence of differential impact of health reform on racial/ethnic differences in hospital admissions

Table 3 | Changes in rates of preventable hospital admissions per 100 000 residents/year in Massachusetts and control states (NY, NJ, PA) before (1 October 2004–30 June 2006) and after (1 January 2008–30 September 2009) healthcare reform according to race and ethnicity

ASCS measures	Massachusetts			Control states			Differences in differences estimates		Adjusted estimated % change (95% CI)†
	Before	After	% change	Before	After	% change	Unadjusted	Adjusted (95% CI)*	
Overall composite									
White	667	647	−3.0	716	680	−5.1	2.1	2.1 (−0.8 to 5.0)	Ref
Black	1713	1744	1.8	2188	2240	2.4	−0.6	−0.5 (−6.0 to 5.3)	−1.9 (−8.5 to 5.1)
Hispanic	1258	1203	−4.4	1126	1024	−9.1	4.7	1.6 (−3.9 to 5.5)	2.0 (−7.5 to 12.4)
Acute composite									
White	285	263	−7.5	277	262	−5.6	−1.9	−1.8 (−5.2 to 1.7)	Ref
Black	496	470	−5.3	482	476	−1.2	−4.0	−4.0 (−12.2 to 5.1)	−1.4 (−12.7 to 11.4)
Hispanic	393	362	−7.8	297	276	−7.3	−0.5	−1.2 (−9.9 to 8.3)	2.0 (−10.3 to 15.7)
Chronic composite									
White	383	384	0.3	440	419	−4.7	5.0	5.0 (1.6 to 8.6)	Ref
Black	1217	1274	4.7	1706	1764	3.4	1.3	1.3 (−4.9 to 7.9)	−3.1 (−9.4 to 3.7)
Hispanic	865	840	−2.8	829	748	−9.7	6.9	2.9 (−3.4 to 9.5)	−0.7 (−9.6 to 12.2)

*Adjusted difference in differences estimates and 95% CI obtained from Poisson regression models adjusted for sex, age, race/ethnicity, county income level, county unemployment rate, quarter, and Health Professions Shortage Area designation.
†For change in racial/ethnic disparities in MA v controls. Expresses change in disparities after reform between black and white people and between Hispanic and white people in ACSC (preventable hospitalization) rates after adjustment for changes in control states.

McCormick et al. (2015)

BMC Public Health

RESEARCH ARTICLE   **Open Access**

Did the English strategy reduce inequalities in health? A difference-in-difference analysis comparing England with three other European countries

Yannan Hu[1], Frank J. van Lenthe[1], Ken Judge[2], Eero Lahelma[3], Giuseppe Costa[4], Rianne de Gelder[1] and Johan P. Mackenbach[1*]

Abstract

Background: Between 1997 and 2010, the English government pursued an ambitious programme to reduce health inequalities, the explicit and sustained commitment of which was historically and internationally unique. Previous evaluations have produced mixed results. None of these evaluations have, however, compared the trends in health inequalities within England with those in other European countries. We carried out an innovative analysis to assess whether changes in trends in health inequalities observed in England after the implementation of its programme, have been more favourable than those in other countries without such a programme.

Methods: Data were obtained from nationally representative surveys carried out in England, Finland, the Netherlands and Italy for years around 1990, 2000 and 2010. A modified difference-in-difference approach was used to assess whether trends in health inequalities in 2000–2010 were more favourable as compared to the period 1990–2000 in England, and the changes in trends in inequalities after 2000 in England were then compared to those in the three comparison countries. Health outcomes were self-assessed health, long-standing health problems, smoking status and obesity. Education was used as indicator of socioeconomic position.

Results: After the implementation of the English strategy, more favourable trends in some health indicators were observed among low-educated people, but trends in health inequalities in 2000–2010 in England were not more favourable than those observed in the period 1990–2000. For most health indicators, changes in trends of health inequalities after 2000 in England were also not significantly different from those seen in the other countries.

Conclusions: In this rigorous analysis comparing trends in health inequalities in England both over time and between countries, we could not detect a favourable effect of the English strategy. Our analysis illustrates the usefulness of a modified difference-in-difference approach for assessing the impact of policies on population-level health inequalities.

Keywords: Health inequality, English strategy, Self-assessed health, Long-standing health problems, Obesity, Smoking, Difference-in-difference analysis, Europe

Abbreviations: BMI, body mass index; OR, odds ratio

What was the English "strategy"?
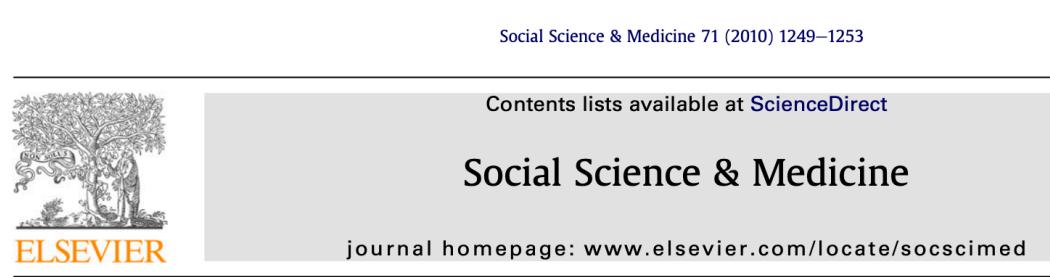
National policies in 1999

Tackling low income, family support policies, tax-reduction and long-term care for the elderly, anti-smoking policies, improving early education

National policies in 2003

Reducing poor health behaviors in manual social groups, improving housing quality, and reducing accidents at home and on the road.

Hu et al. (2016)

These policies showed little evidence for inequality reduction *in England* ⋯

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Has the English strategy to reduce health inequalities failed?

Johan P. Mackenbach*

Department of Public Health, Erasmus MC, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands

⋯even if there is no more reduction in health inequalities after the implementation of the strategy than before, the changes in trends in England could still be more favourable than those in other European countries that have done less to reduce health inequalities. - Hu et al. (2016)

Mackenbach (2010)

Data on self-reported health, smoking, obesity

Table 1 Countries included in the analysis and sources of data

Country	Survey year	Survey names
England	1991–1992; 2000; 2010	Health Survey for England
Finland	1989; 1999; 2009	Health Behaviour and Health
The Netherlands	1990	Ongoing Survey of Living Conditions (DLO)
	2000; 2009	Permanent Survey of Living Conditions (POLS)
Italy	1990	Multipurpose Family Survey
	2000	Health and Health Care Utilization
	2010	Multipurpose Family Survey-Aspects of daily living

The Finnish data used in this study are the data combined from the two Finish studies: "Health behaviour and health among Finnish adult population (AVTK)", which includes respondents who are 15–64 years old, and "Health behaviour and health among the Finnish elderly (EVTK)", which includes respondents who are older than 64 years

Hu et al. (2016)

For comparison we selected countries that were in a similar stage of awareness of health inequalities, but that had not implemented a national strategy to tackle health inequalities.

- Treatment: 2000s vs. 1990s
- Comparison: England vs. Others
- SEP: Low vs. High Education
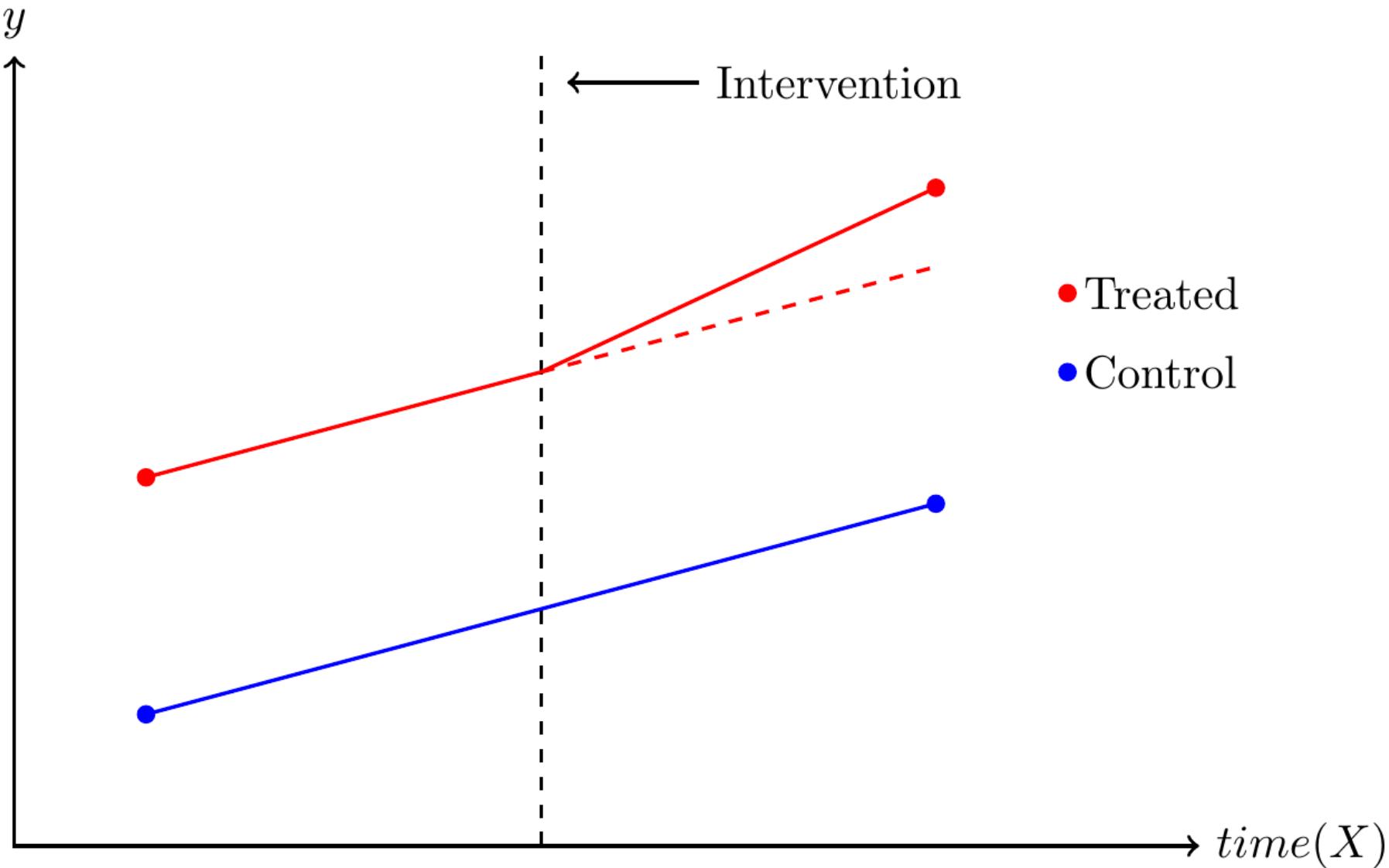
In the last and our main step, we added each of the comparison countries separately to the analysis of the English data, following the idea of "difference-in-differences analysis". Our aim was to investigate whether the changes in trends in health inequalities between 1990–2000 and 2000–2010 were more favourable in England than those in the three comparison countries.

Hu et al. (2016)

Regression-based approach:
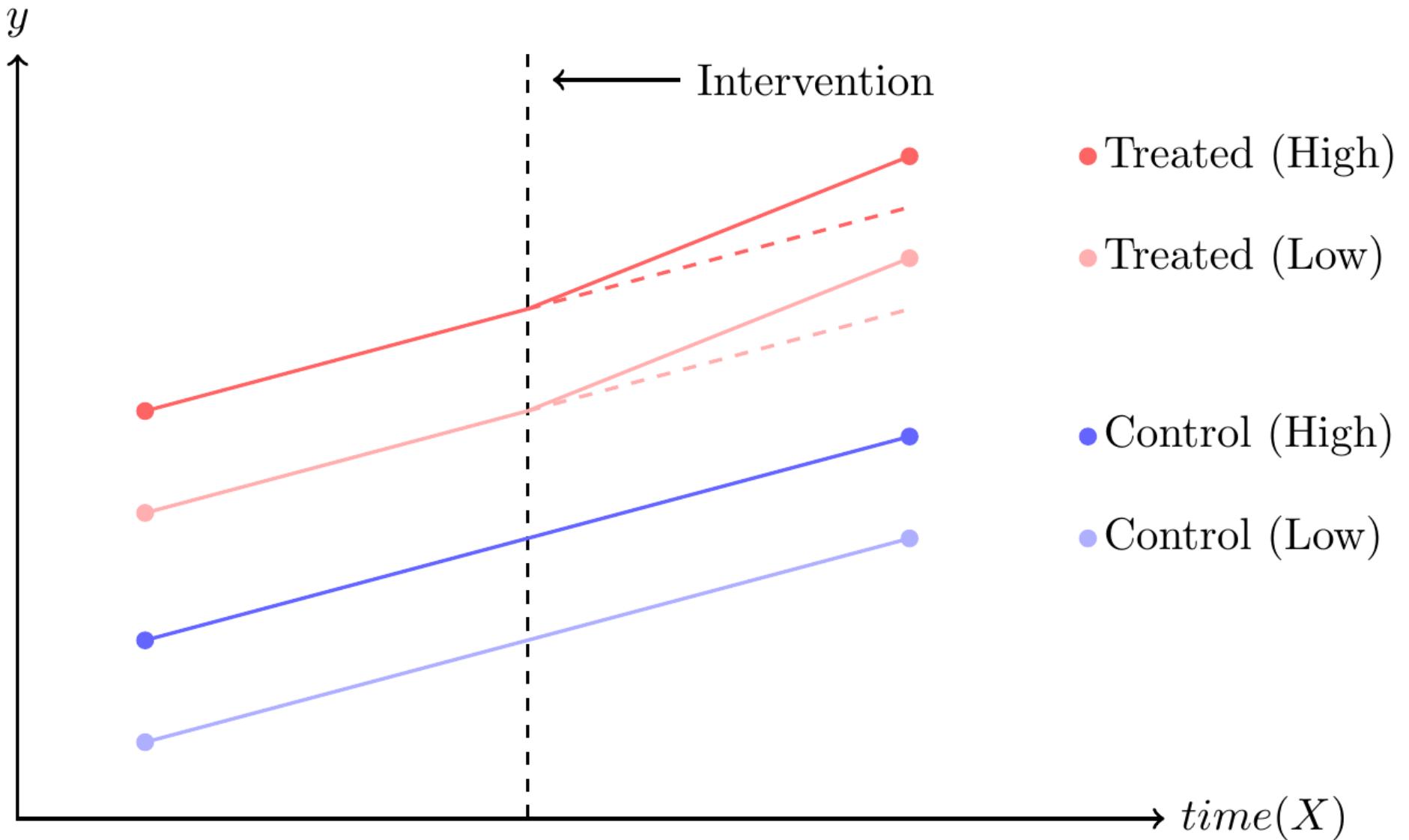
The model can be written as:

$$
\begin{aligned}
outcome_{istj} = (&\beta_0 + \beta_1 endyear_{tj} + \beta_2 policyperiod_{sj} \\
&+ \beta_3 endyear_{tj} * policyperiod_{sj} \\
&+ \beta_4 ledu_{istj} + \beta_5 ledu_{istj} * endyear_{tj} \\
&+ \beta_6 ledu_{istj} * policyperiod_{sj} \\
&+ \beta_7 ledu_{istj} * endyear_{tj} * policyperiod_{sj} \\
&+ X_{istj}) + (\beta_0' + \beta_1' endyear_{tj} + \beta_2' policyperiod_{sj} \\
&+ \beta_3' endyear_{tj} * policyperiod_{sj} \\
&+ \beta_4' ledu_{istj} + \beta_5' ledu_{istj} * endyear_{tj} \\
&+ \beta_6' ledu_{istj} * policyperiod_{sj} \\
&+ \beta_7' ledu_{istj} * endyear_{tj} * policyperiod_{sj} \\
&+ X_{istj}) * england_j
\end{aligned}
$$
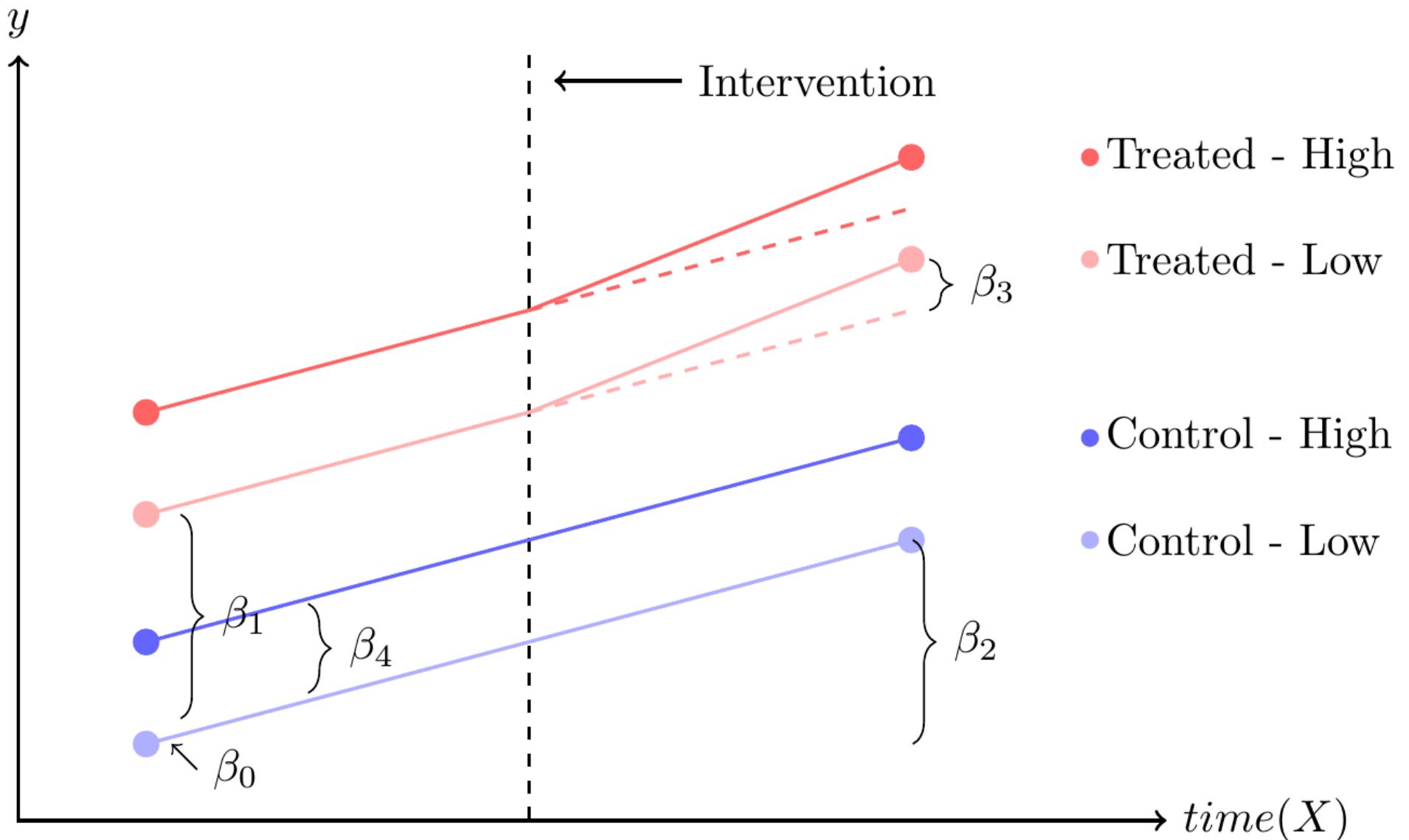
Our basic setup was for group *averages*
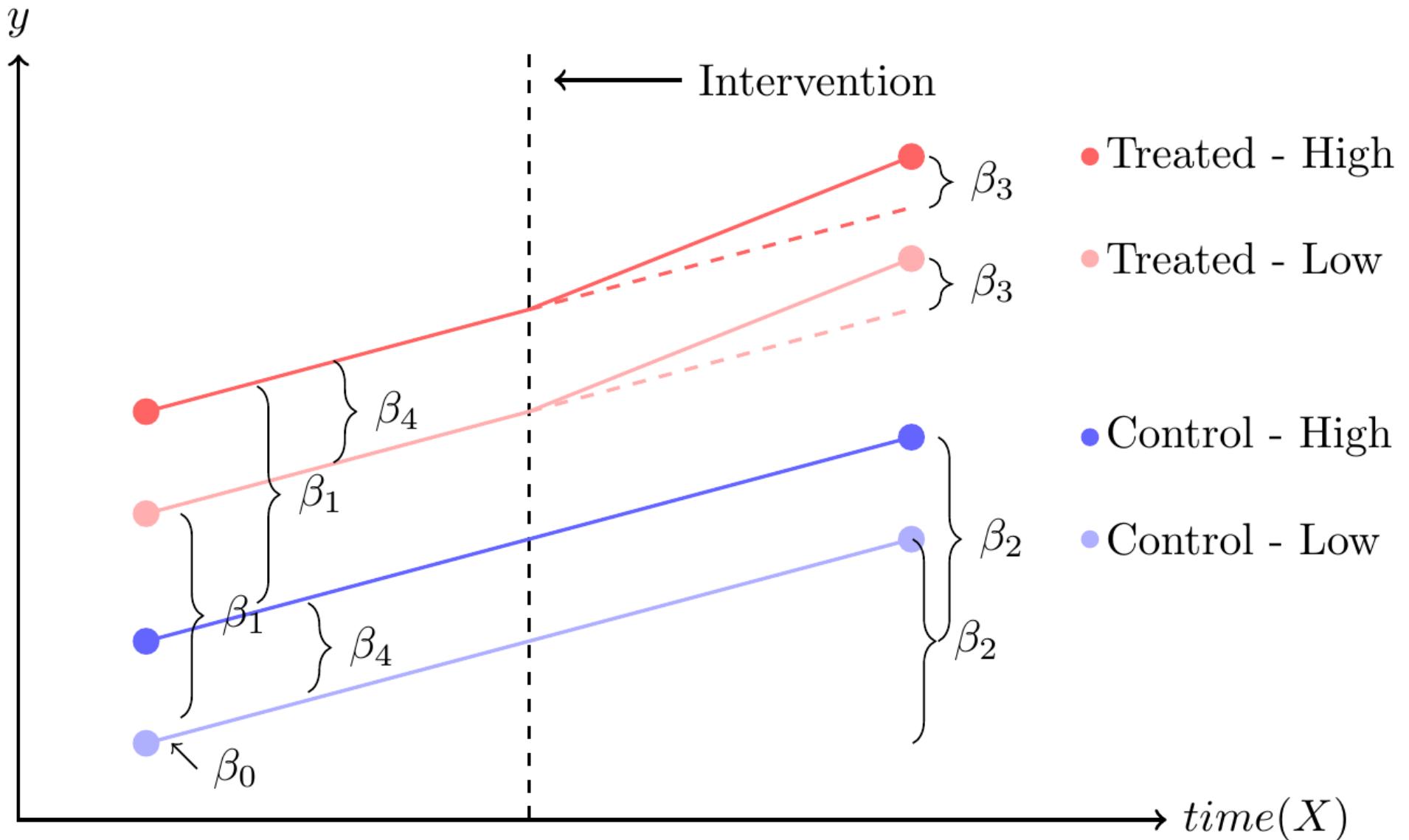
Let's add two social groups per unit

$$Y = \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 Treat * Post + \beta_4 High$$
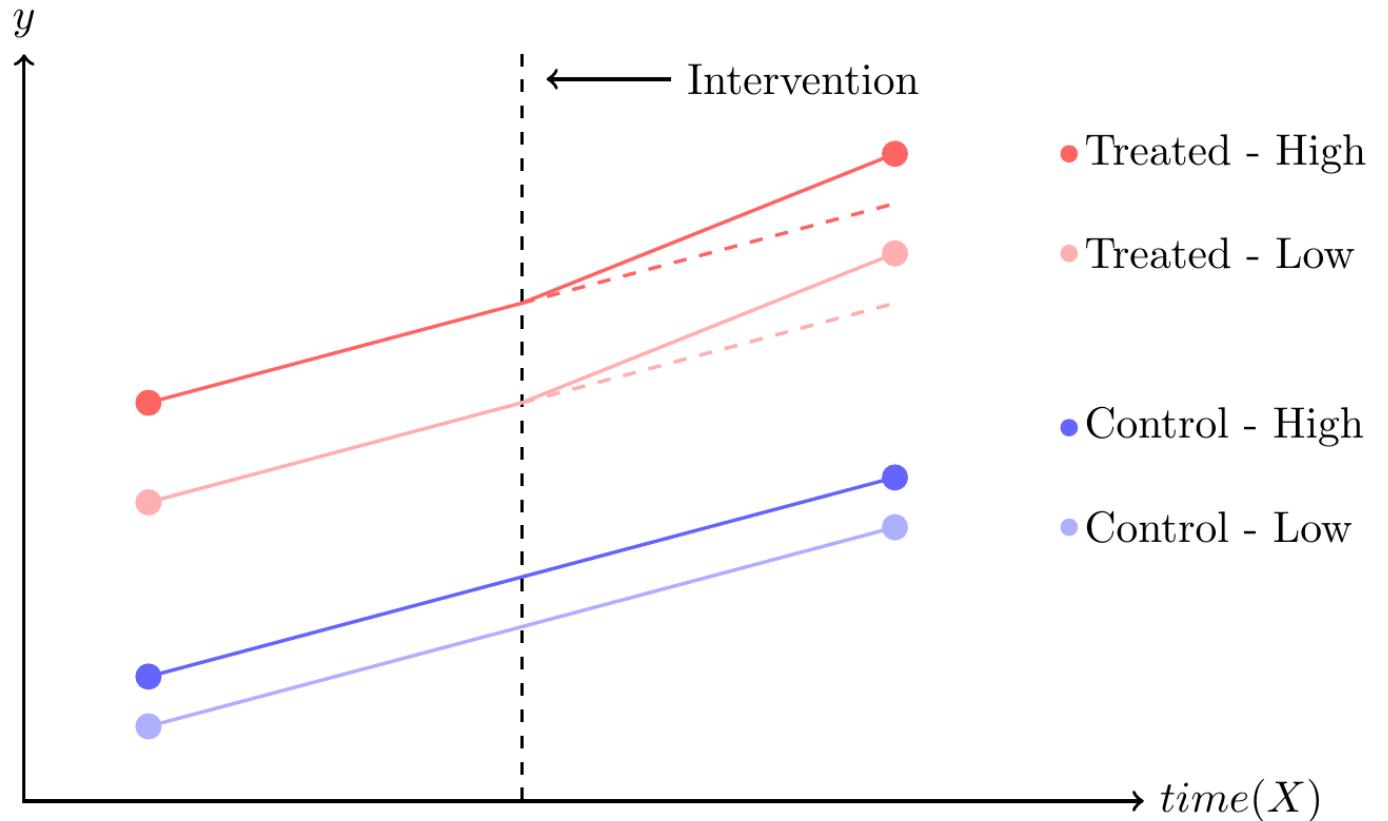
$$Y = \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 Treat * Post + \beta_4 High$$
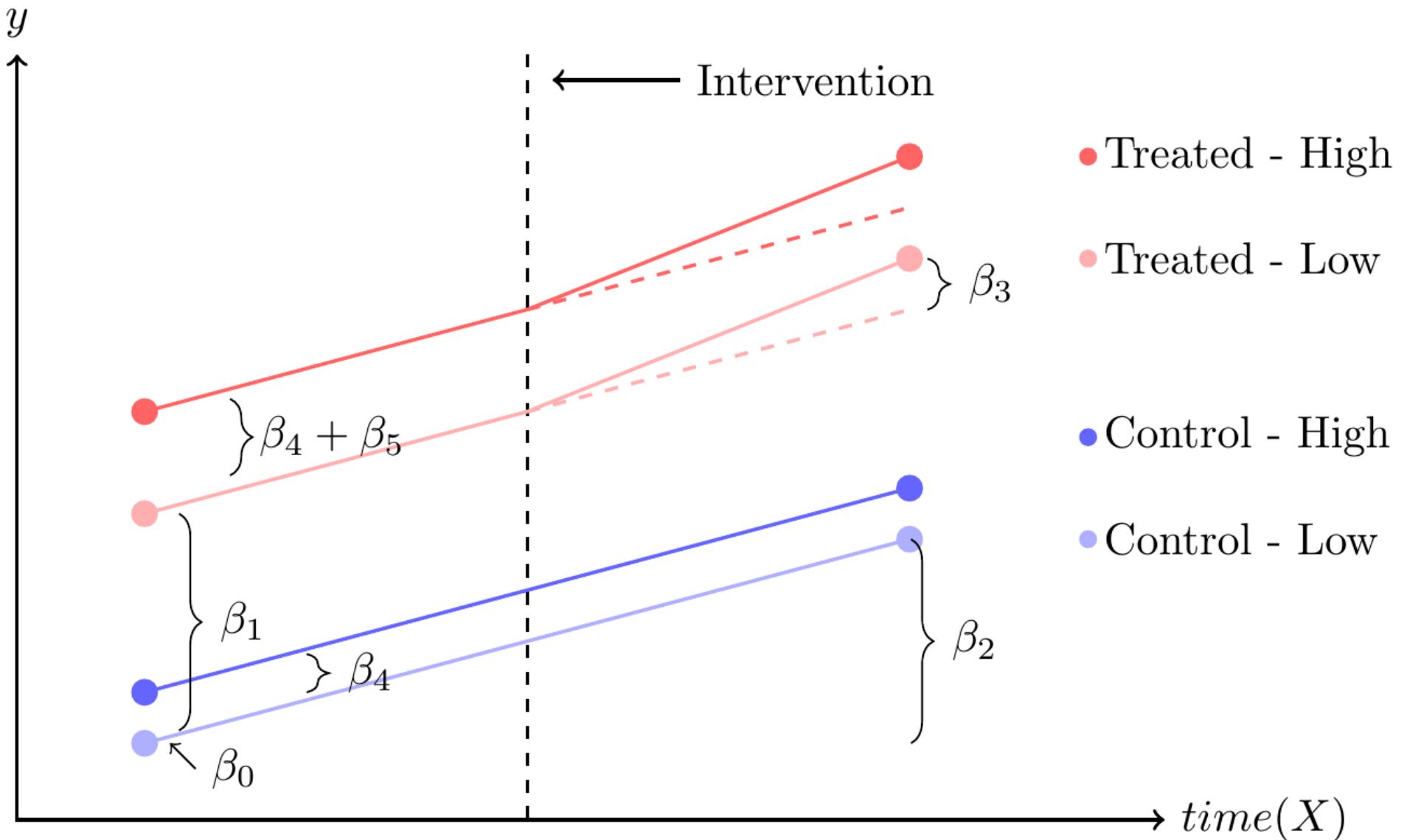
Allowing for SEP-group heterogeneity

- Socioeconomic differences may be different magnitude in treated vs. control areas.

- Better resources, more advocacy, different demographics, etc.
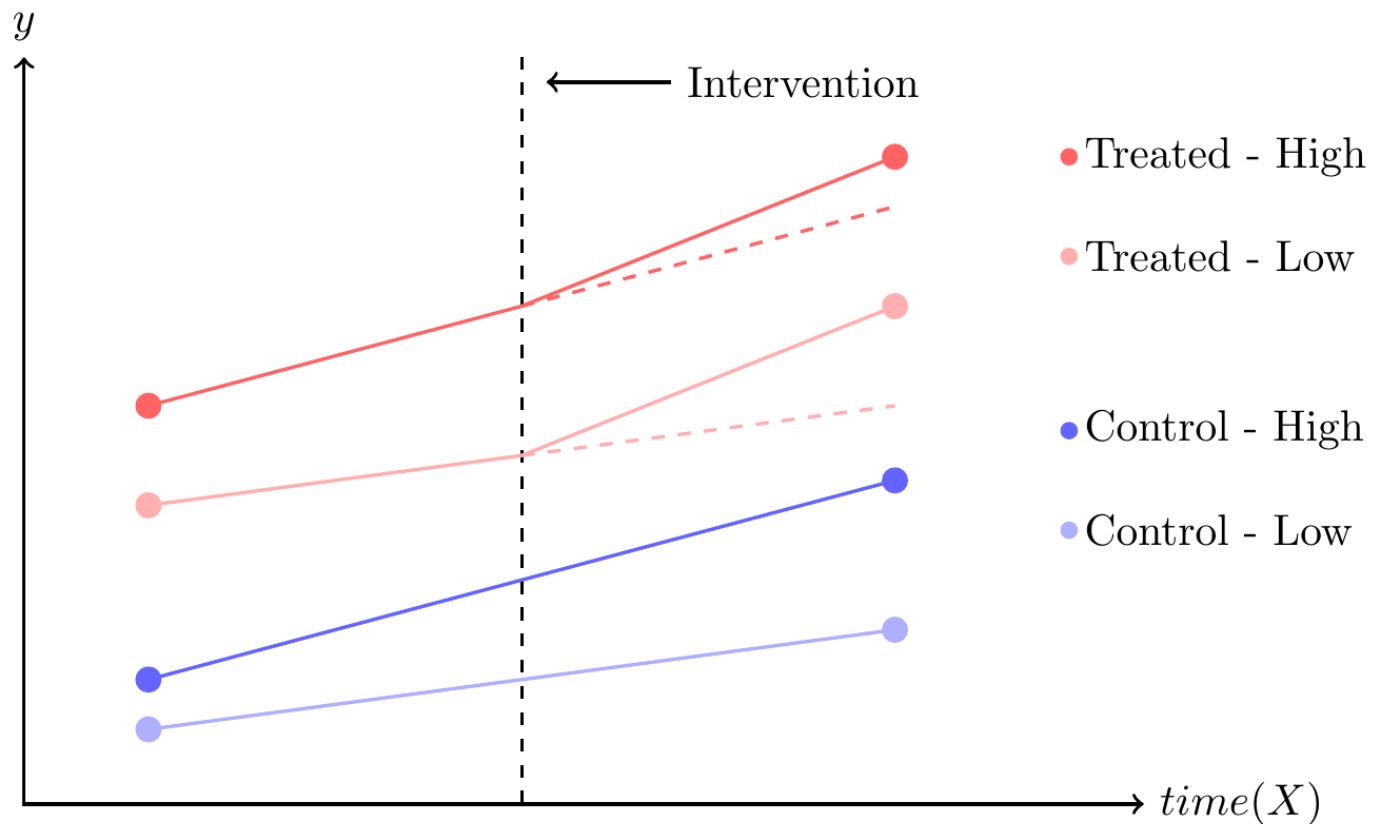
$$Y = \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 Treat*Post + \beta_4 High + \beta_5 Treat*High$$
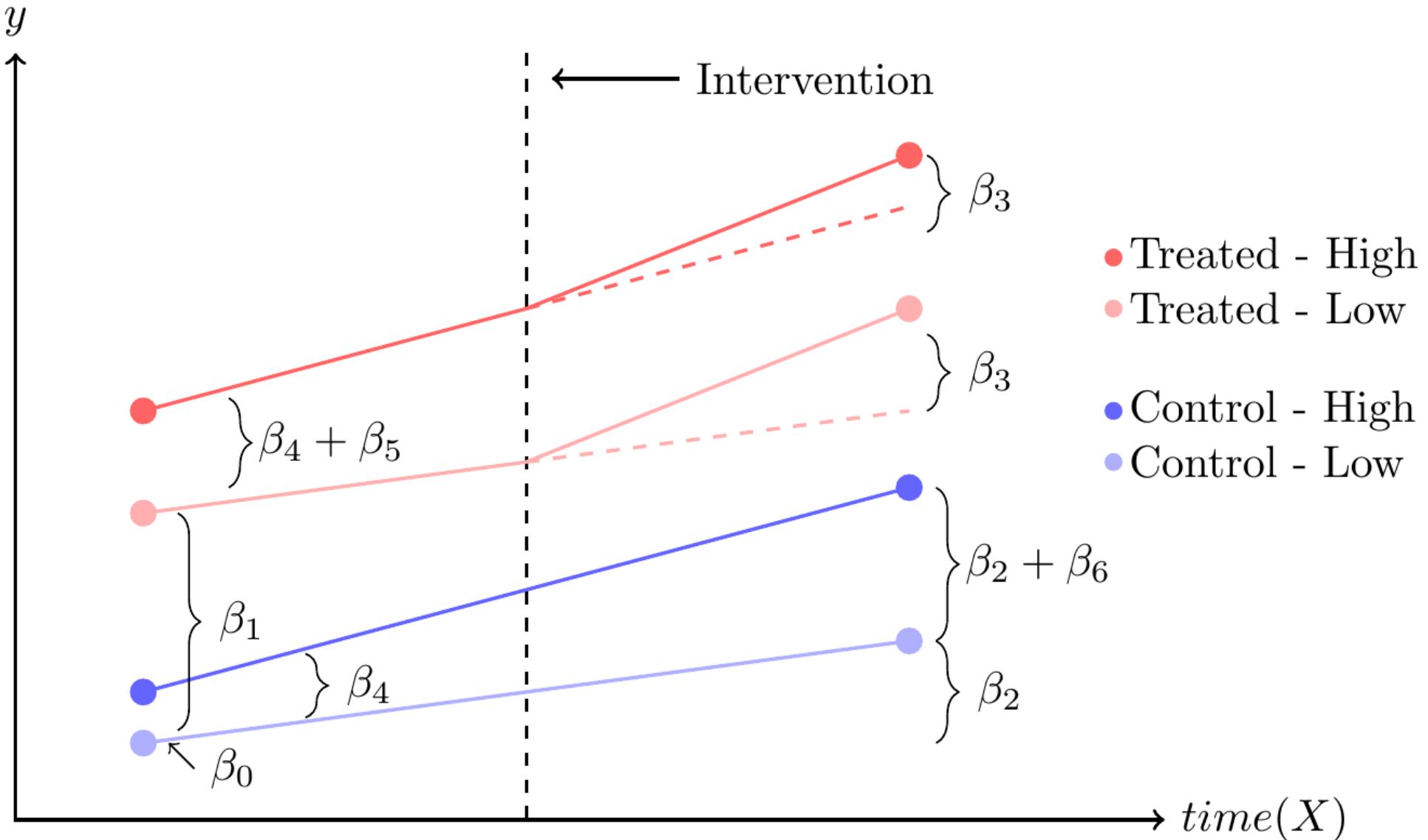
Allowing for SEP-time heterogeneity
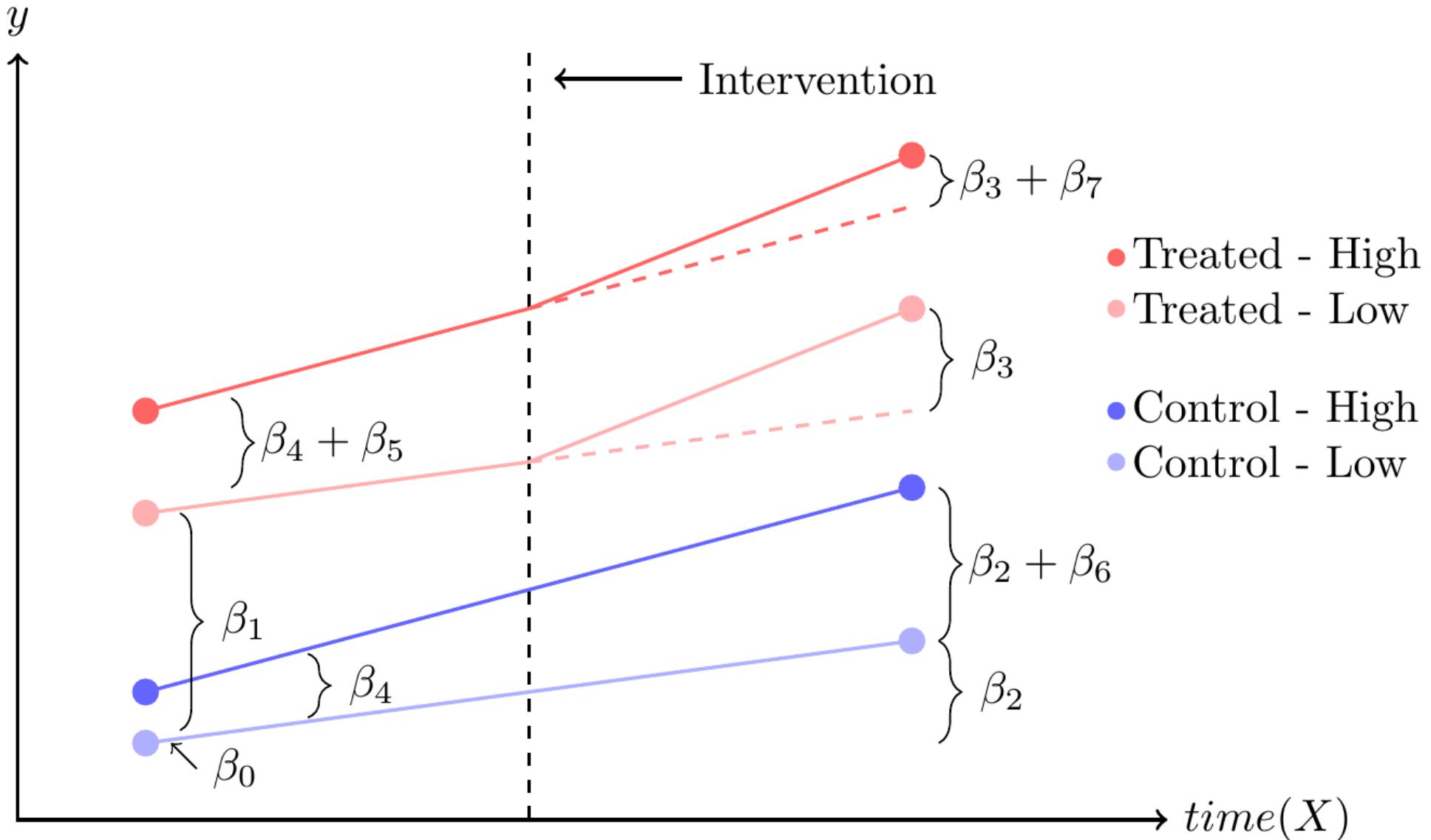
- Secular trends may be changing differentially by social group in *all* areas.

- Different baseline health, health behaviors, access to resources, etc.

$$Y = \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 Treat * Post + \beta_4 High + \beta_5 Treat * High + \beta_6 Post * High$$

$Y = \beta_0 + \beta_1 Treat + \beta_2 Post + \beta_3 Treat * Post + \beta_4 High + \beta_5 Treat * High + \beta_6 Post * High + \beta_7 Treat * Post * High$

|	Odds ratios (logistic)			
	Less-than-good self-assessed health	Long-standing health problems	Smoker	Obesity
1. Two-way interaction parameter estimates[a]				
England (low-edu)	0.76***	0.78***	0.82**	0.97
	(0.064)	(0.065)	(0.073)	(0.097)
2. Three-way interaction parameter estimates[b]				
England	1.22	0.95	1.19	1.25
	(0.197)	(0.125)	(0.182)	(0.213)
Finland	0.78	–	1.28	1.90*
	(0.173)	–	(0.308)	(0.652)
The Netherlands	1.18	1.16	1.00	–
	(0.221)	(0.181)	(0.165)	–
Italy	–	–	0.97	0.76*
	–	–	(0.072)	(0.121)
3. Four-way interaction parameter estimates[c]				
England vs Finland	1.57	–	0.93	0.66
	(0.433)	–	(0.267)	(0.253)
England vs the Netherlands	1.04	0.82	1.20	–
	(0.257)	(0.167)	(0.270)	–
England vs Italy	–	–	1.23	1.64**
	–	–	(0.209)	(0.383)

Results showed that changes in trends of inequalities after 2000 were not statistically significantly different between England and any of the other countries, with the single exception of obesity for which the change was less favourable in England than in Italy (OR = 1.64, p < 0.05).

	Odds ratios (logistic)			
	Less-than-good self-assessed health	Long-standing health problems	Smoker	Obesity
1. Two-way interaction parameter estimates[a]				
England (low-edu)	0.76***	0.78***	0.82**	0.97
	(0.064)	(0.065)	(0.073)	(0.097)
2. Three-way interaction parameter estimates[b]				
England	1.22	0.95	1.19	1.25
	(0.197)	(0.125)	(0.182)	(0.213)
Finland	0.78	–	1.28	1.90*
	(0.173)	–	(0.308)	(0.652)
The Netherlands	1.18	1.16	1.00	–
	(0.221)	(0.181)	(0.165)	–
Italy	–	–	0.97	0.76*
	–	–	(0.072)	(0.121)
3. Four-way interaction parameter estimates[c]				
England vs Finland	1.57	–	0.93	0.66
	(0.433)	–	(0.267)	(0.253)
England vs the Netherlands	1.04	0.82	1.20	–
	(0.257)	(0.167)	(0.270)	–
England vs Italy	–	–	1.23	1.64**
	–	–	(0.209)	(0.383)

The interpretation of the interaction terms in difference-in-differences logistic models is essentially similar to that in the more common linear models, except that they indicate the relative change of the odds of the health outcome in the treatment group relative to that in the control group, instead of the absolute change of the rate of the health outcome in the treatment group minus that in the control group

- Parallel trends assumption is scale dependent.

- Can't have it both ways.

Differences in levels

Differences in logs
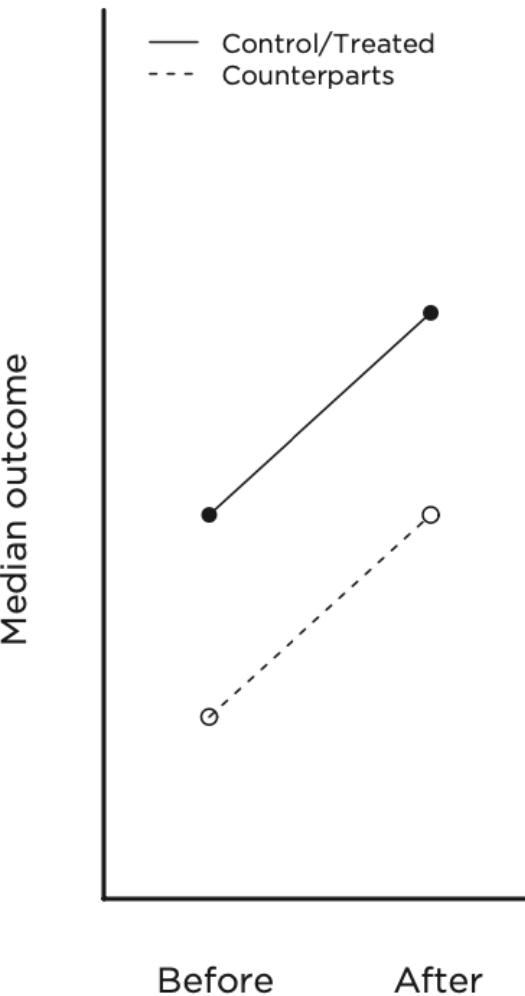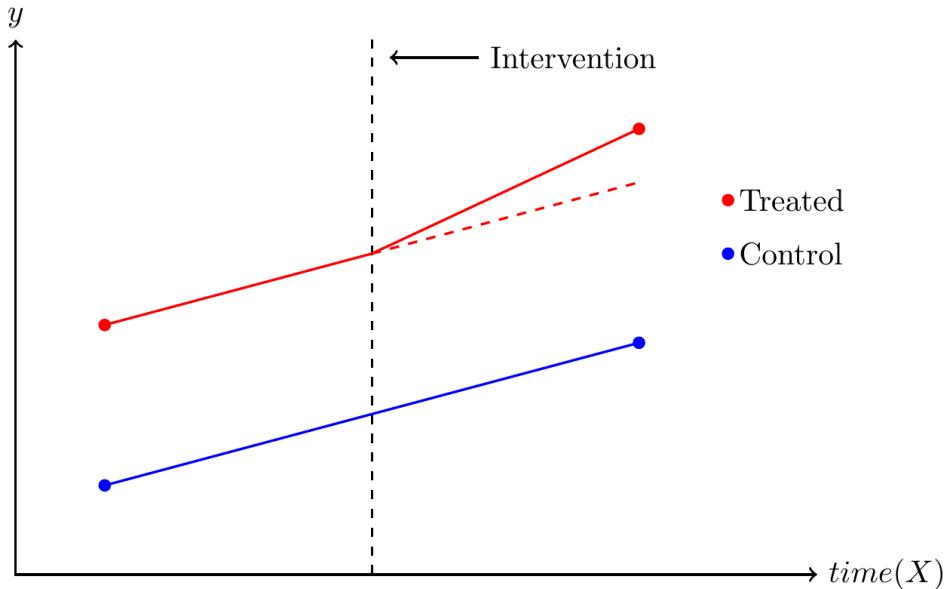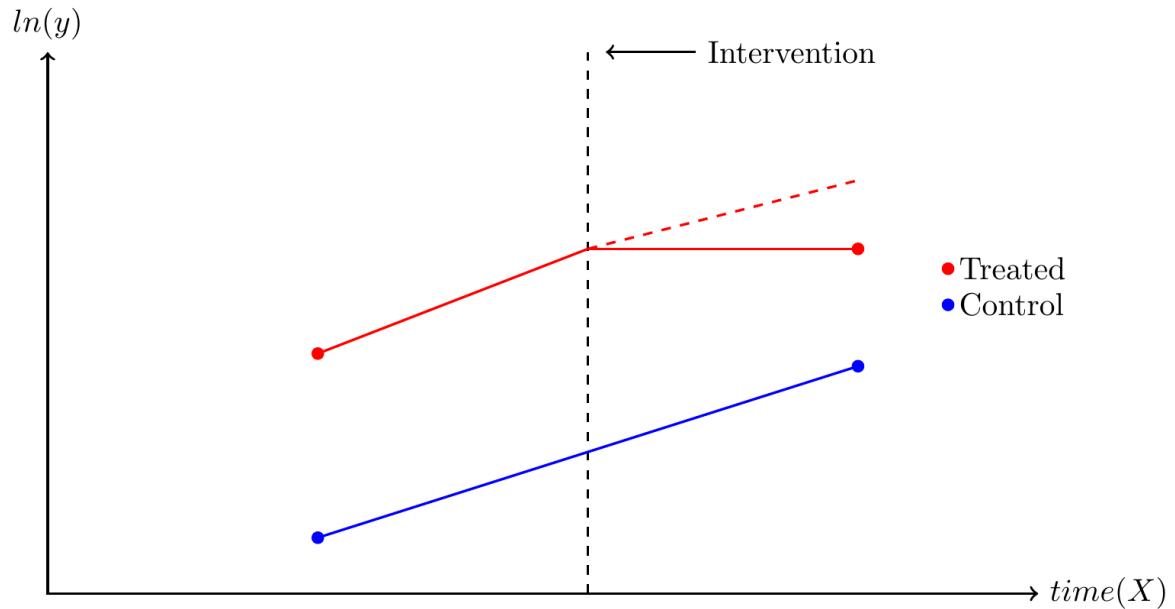
Figure from Rosenbaum (2017); see also Roth and Sant'Anna (2023)

Implications for counterfactuals

The change in the treated group in both graphs is identical (from 1.4 to 2.5).
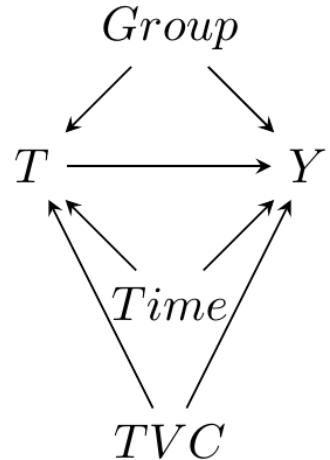
Parallel trends in levels ($\Delta 0.8$) consistent with positive impact of treatment.

Parallel trends in logs ($\times 1.3$) consistent with negative impact of treatment.

4. Extensions

Extending the basic 2x2 DD

Note that our basic regression model assumes the only time-varying factor is the policy:

$$Y_{gt} = \alpha + \gamma_g + \tau_t + \textcolor{red}{\delta^{DD}} p_{gt} + \varepsilon_{gt}$$

What if there are confounders of the decision to change the policy?

We may have omitted important factors that:

- differ by treatment status.
- affect the outcome.
- are time-varying, but not affected by the treatment (TVC).

The literature on covariates is evolving rapidly. See Caetano et al. (2022) for more details.

Adding time-varying covariates

- Suppose the policy is a soft drink tax and the outcome calories consumed (linear).

- We might worry that *changes in* the density of fast food restaurants could be a common cause of both. Now add measured time-varying confounders:

$$Y_{gt} = \alpha + \gamma_g + \tau_t + \textcolor{red}{\delta^{DD}} p_{gt} + \zeta Z_{gt} + \varepsilon_{gt}$$
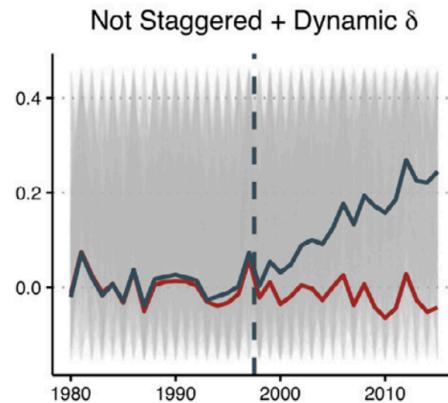
- where ζZ_{gt} is a vector of other controls at the cluster level.

- Important especially if you think other policies may have been implemented simultaneously with treatment.

- Now, conditional on FEs and ζZ_{gt}, we assume that the timing of the change in policy is as good as random.

Extensions to non-binary treatments

DD design can also handle treatments, policies, or exposures that are not dichotomous.

- E.g., changes in minimum wage levels (varying "treatment" intensity)
  - Estimate varying levels of increase ($2 increase vs. $4)
- "Sin" taxes (e.g., alcohol or cigarettes).
  - differential changes in excise taxes (smaller vs. larger).
- "Weaker" vs. "Stronger" policies
  - texting while driving (primary vs. secondary offense)
 - thresholds for blood alcohol limits (0.15 vs. 0.10 vs. 0.08).

Dynamic Effects

Not Staggered + Dynamic δ
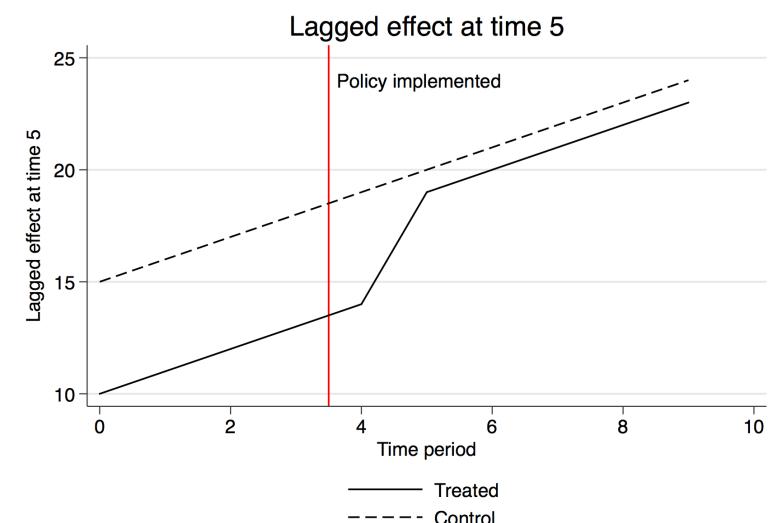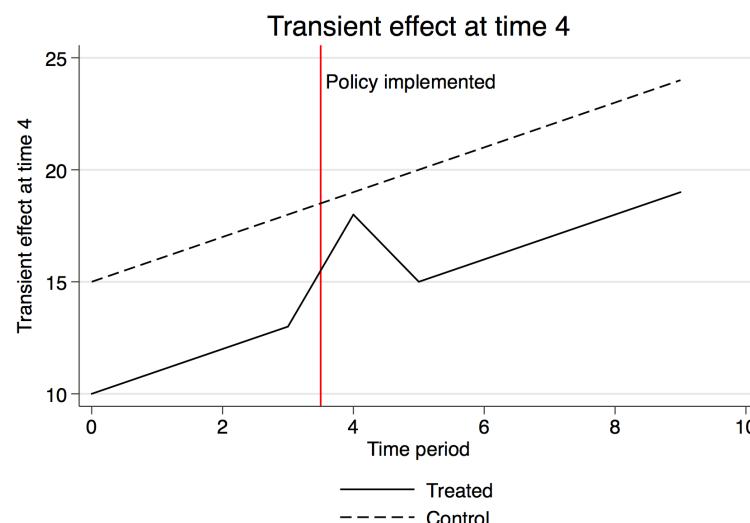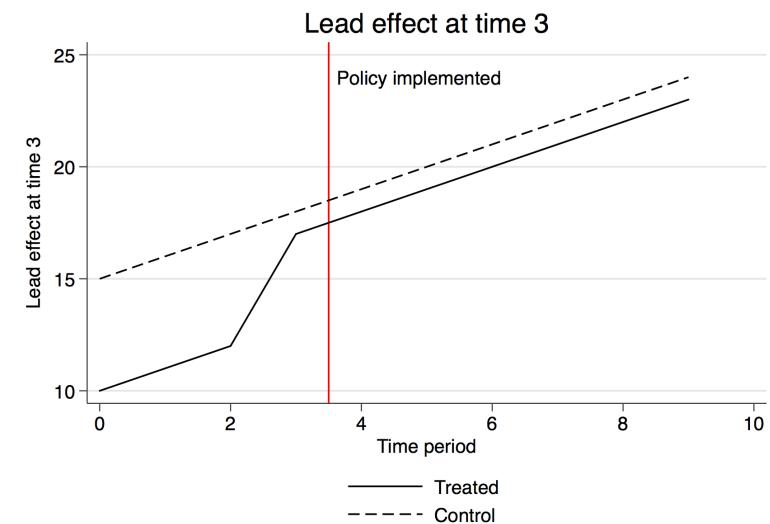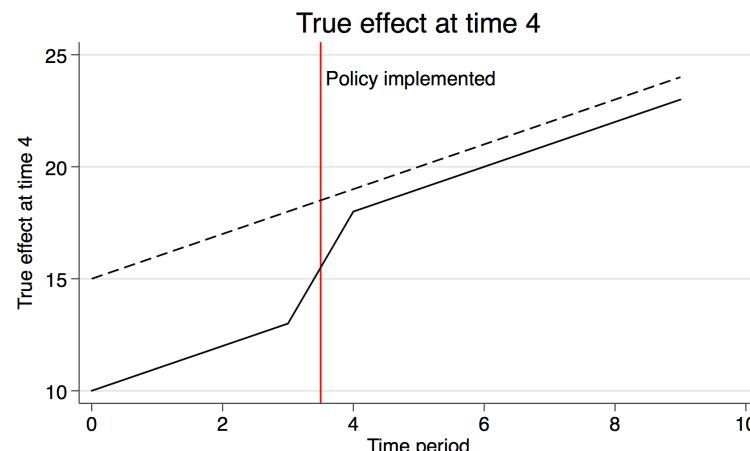
- Basic DD estimates the *average ATT* over the entire post-intervention period.

- May average over important variations in how the treatment evolves over time.

- Was the impact immediate? Transient? Sustained over time?

- Can extend the basic model to allow for heterogeneity over time.

Figure from Baker et al. (2022)

Hypothetical dynamic treatment effect scenarios

What about staggered treatments?

- Different groups adopt treatments at different times.

- Creates many 2x2 DDs.
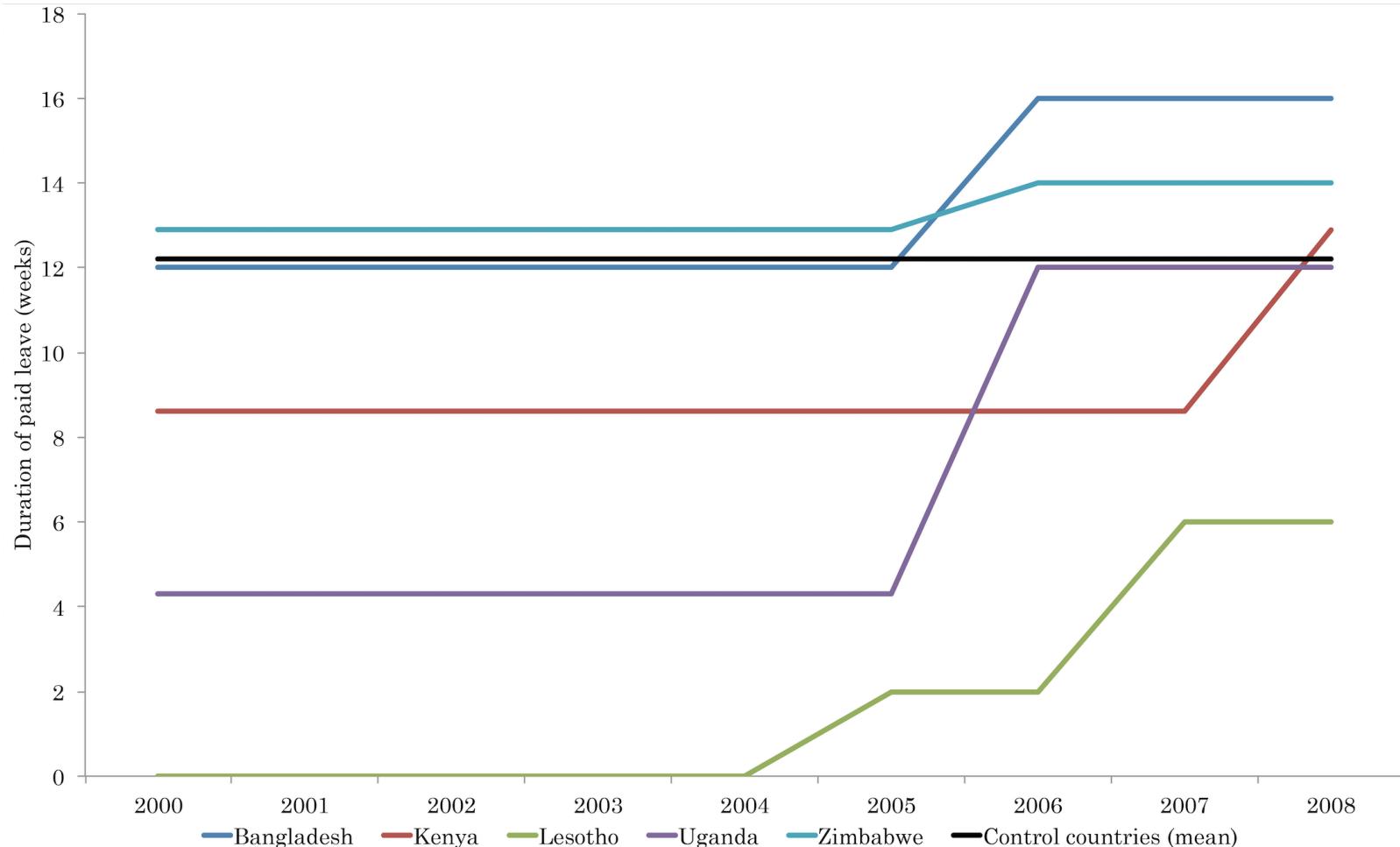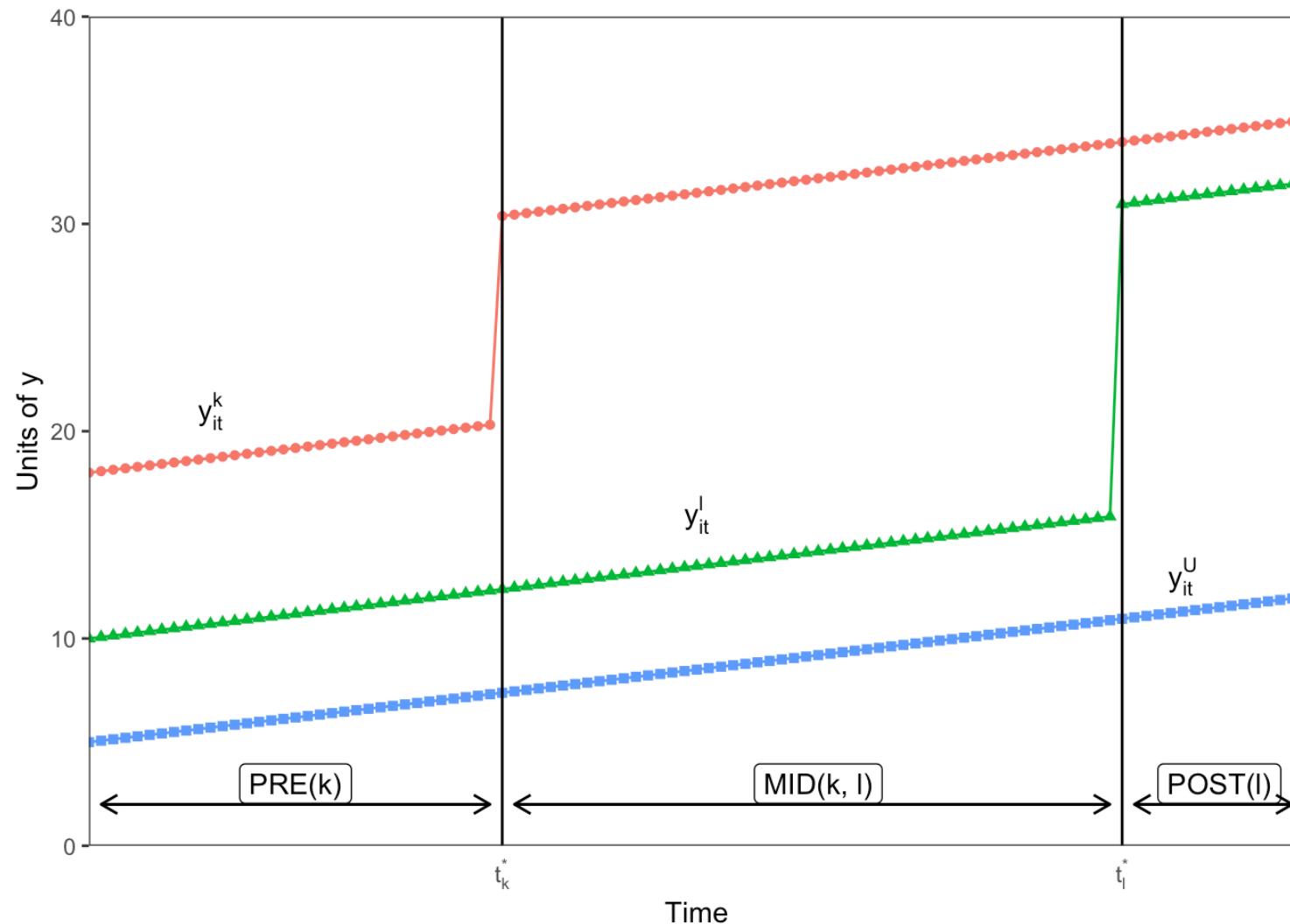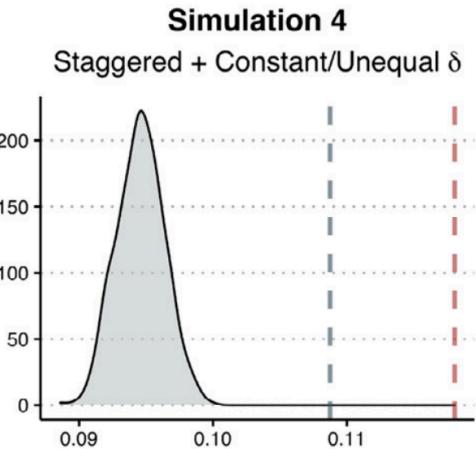
Figure from Nandi et al. ([2016](#))

1. Early-adopters (k) vs. never treated (U)

2. Later-adopters (l) vs. never treated (U).

3. Early (k) vs. later (l) adopters.

4. Later (l) vs. earlier (k) adopters.

Graph from https://andrewcbaker.netlify.app/2019/09/25/difference-in-differences-methodology/

What is the problem?

Simulation 4
Staggered + Constant/Unequal δ

- Using earlier treated groups as controls only 'works' if the treatment effects are:

  - Homogeneous across groups at a given time; and

  - Homogeneous over time (no dynamic effects).

- This adds any changes in treatment effects in the early group, which get subtracted from the DD estimate.

- Can lead to β^{DD} that is a poor summary of group-specific effects if there is heterogeneity.

Figure from Baker et al. (2022)

What are potential solutions?

- All basically involve not allowing early treated groups to serve as controls later.

Callaway and Sant'Anna

Use non-parametric group-time ATTs (+ covariates).

Abraham and Sun

Use saturated fixed effects to ensure that prior treated units are not used as controls

Cengiz, Dube, Lindner, and Zipperer

Create state-event-specific panel datasets and calculate event-specific estimates using separate regressions for each state-event.

Many new papers on this, including Callaway and Sant'Anna (2021), Goodman-Bacon (2021), Cengiz et al. (2019) Sun and Abraham (2021).

Key Takeaways

- DD compares *changes* in outcomes in a treated group to a control group.

- Controls for time-invariant unobserved group factors *and* common trends in outcomes.

- Requires good qualitative knowledge about *why* the treated group became treated.

- Core assumption is parallel trends, unverifiable but not impossible to investigate.

- Can be extended to address inequalities, but stronger assumptions needed.

- Strong designs like DD can help reduce the "evidence gap".

References

Baker, A.C., Larcker, D.F., Wang, C.C.Y., 2022. Journal of Financial Economics 144, 370–395.

Caetano, C., Callaway, B., Payne, S., Rodrigues, H.S., 2022. Difference in Differences with Time-Varying Covariates.

Callaway, B., Sant'Anna, P.H.C., 2021. Journal of Econometrics, Themed Issue: Treatment Effect 1 225, 200–230.

Cengiz, D., Dube, A., Lindner, A., Zipperer, B., 2019. The Quarterly Journal of Economics 134, 1405–1454.

Freedman, D.A., 1991. Sociological Methodology 21, 291.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., Vermeersch, C.M., 2016. Impact evaluation in practice. World Bank
        Publications.

Goodman-Bacon, A., 2021. Journal of Econometrics, Themed Issue: Treatment Effect 1 225, 254–277.

Hu, Y., van Lenthe, F.J., Judge, K., Lahelma, E., Costa, G., de Gelder, R., Mackenbach, J.P., 2016. BMC Public Health 16, 865.

Mackenbach, J.P., 2010. Social Science & Medicine 71, 1249–1253.

McCormick, D., Hanchate, A.D., Lasser, K.E., Manze, M.G., Lin, M., Chu, C., Kressin, N.R., 2015. BMJ 350, h1480–h1480.

Nandi, A., Hajizadeh, M., Harper, S., Koski, A., Strumpf, E.C., Heymann, J., 2016. PLoS Med 13, e1001985.

Petticrew, M., 2007. The European Journal of Public Health 17, 411–413.

Rosenbaum, P.R., 2017. Observation and experiment: An introduction to causal inference. Harvard university press, Cambridge,
        Massachusetts.

Roth, J., Sant'Anna, P.H.C., 2023. Econometrica 91, 737–747.

Snow, J., Frost, W.H., Richardson, B.W., 1936. Snow on cholera: Being a reprint of two papers. Commonwealth Fund, New York.

Sun, L., Abraham, S., 2021. Journal of Econometrics, Themed Issue: Treatment Effect 1 225, 175–199.

Vilhelmsson, A., Östergren, P.-O., 2018. PLOS ONE 13, e0195774.