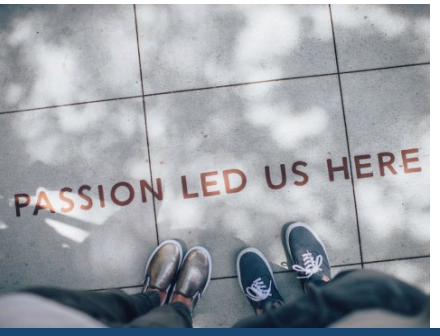


Lab 9: Evaluating Anti-Smoking Policy using Synthetic Control Methods

Gregory Bruich, Ph.D. and Ljubica Ristovska

Note: these slides have been posted on the course website on Canvas; feel free to download to take notes

Spring 2020

HARVARD UNIVERSITY

OPPORTUNITY INSIGHTS

Lab 9: Evaluating Anti-Smoking Policies using Synthetic Control Methods

- In lab 8, we used differences in differences to estimate causal effects by comparing the change over time in an outcome of interest for the treatment group relative to a control group

- *Parallel trends assumption:* changes in the outcome over time *would have* been equal in the treatment vs. control group in the absence of the policy change

- The choice of the control group is therefore crucial, but often *ad hoc*

- In today's lab, we will introduce the **Synthetic Control Method** as a data-driven way of choosing the control group

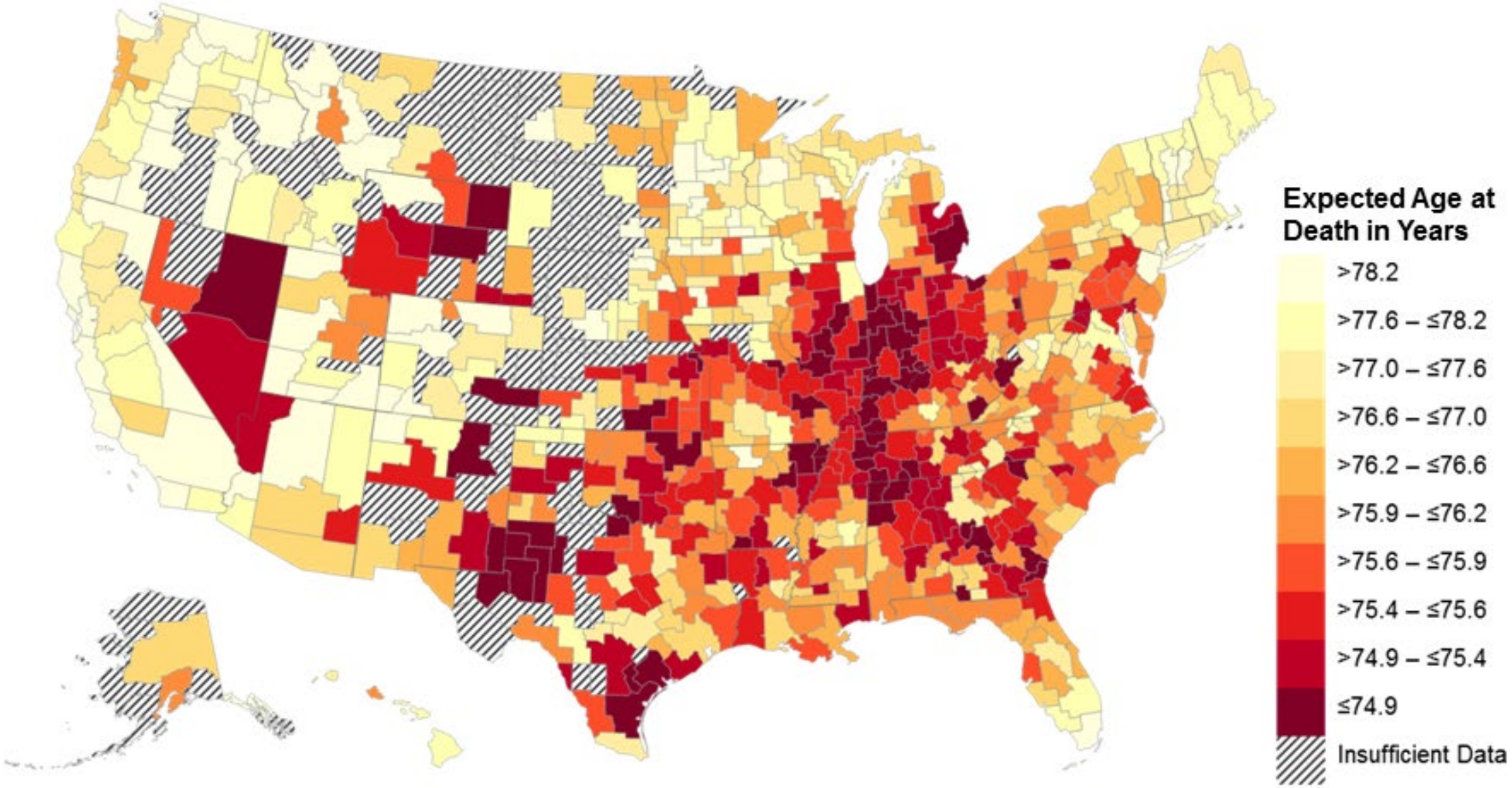
- **Application:** Evaluating the effect of a large-scale tobacco control program in California on smoking (*Abadie, Diamond, and Hainmueller 2010*)

Application

Policies to Improve Life Expectancy

Race-Adjusted Expected Age at Death for 40-Year-Old Men

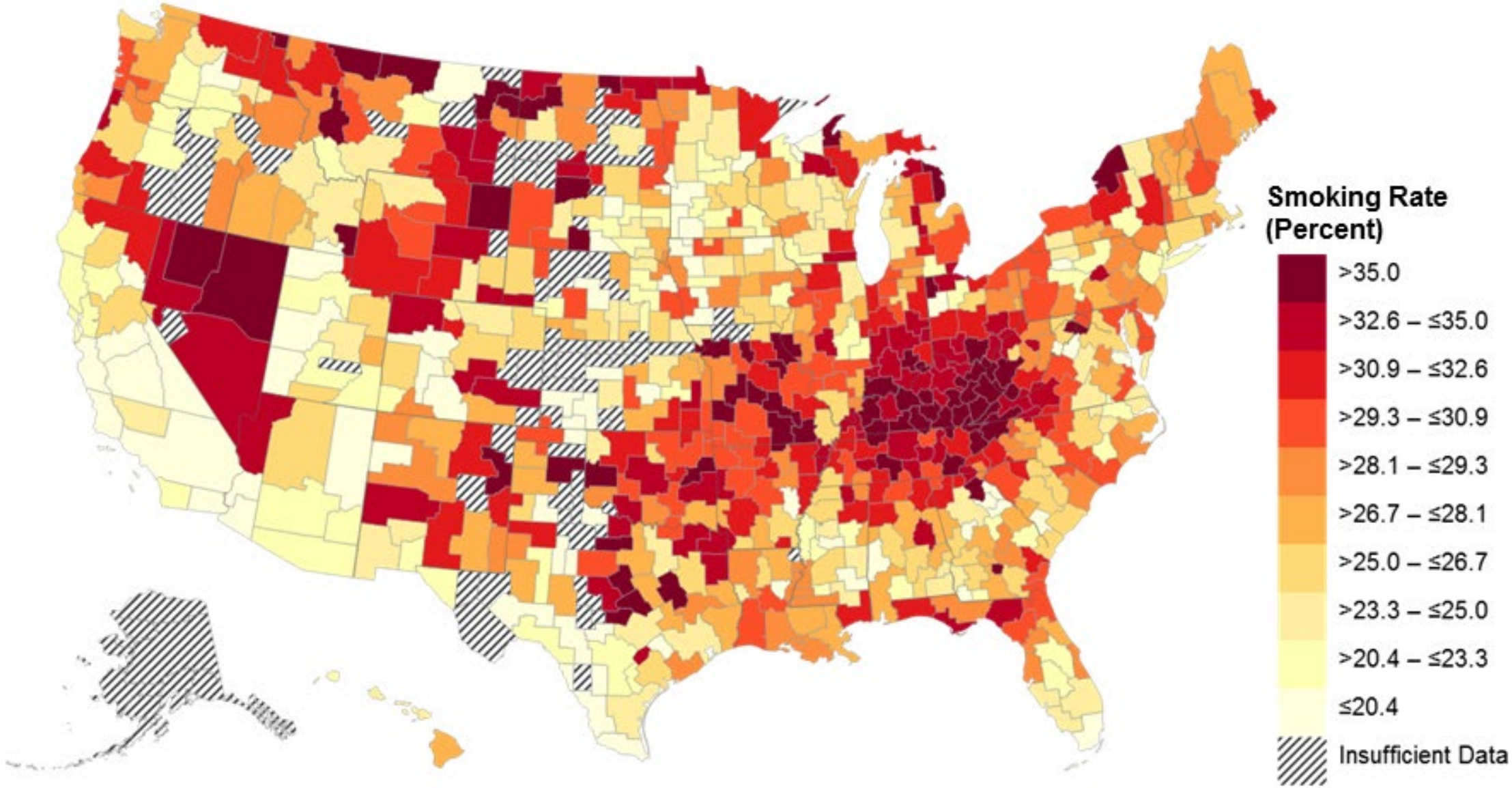
Bottom Quartile of U.S. Income Distribution



Note: Lighter Colors Represent Areas with Higher Life Expectancy

Smoking Rates by Commuting Zone

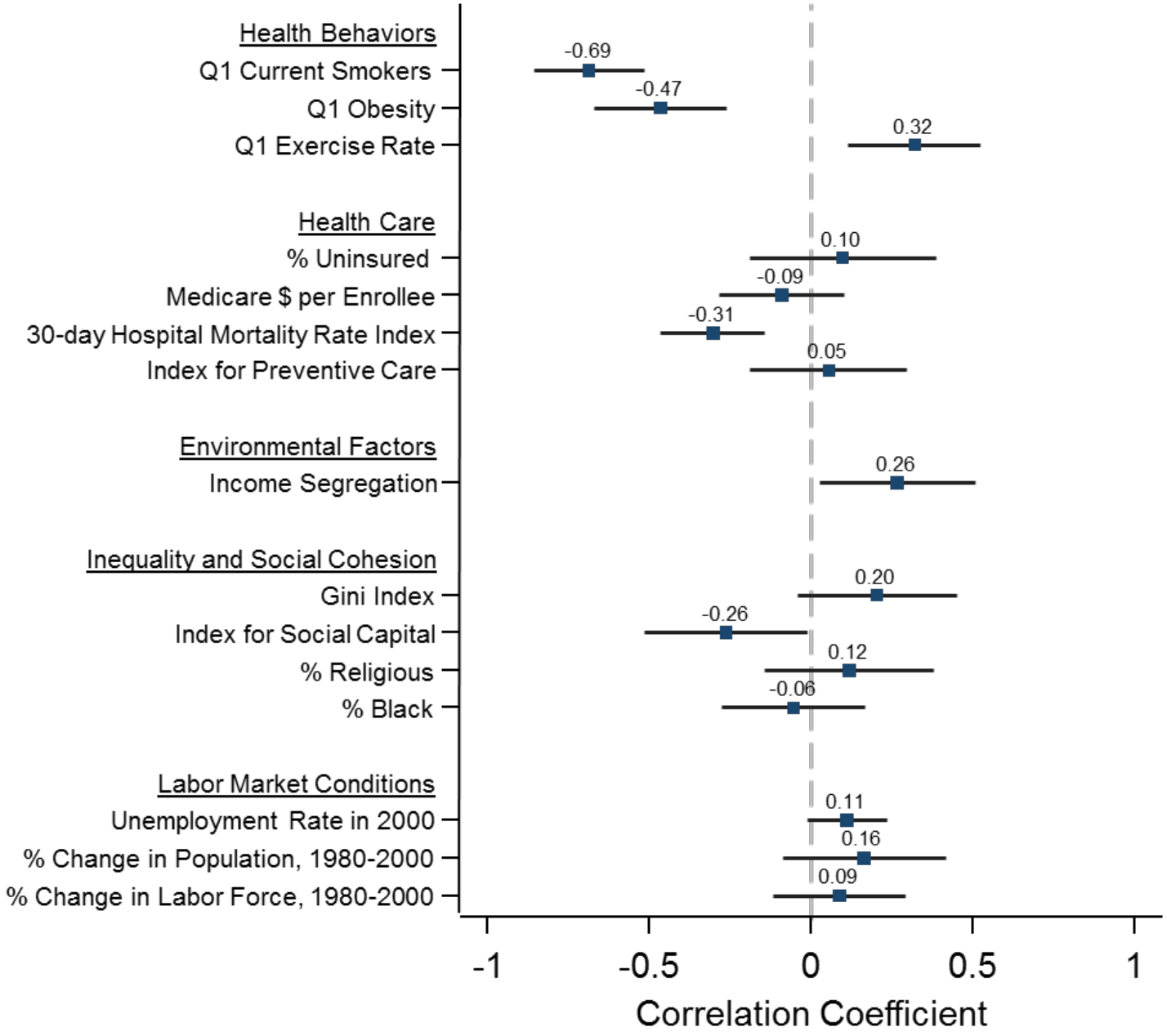
Bottom Quartile of U.S. Income Distribution



Note: Lighter Colors Represent Areas Lower Smoking Rates

Correlations of Expected Age at Death with Health and Social Factors

For Individuals in Bottom Quartile of Income Distribution

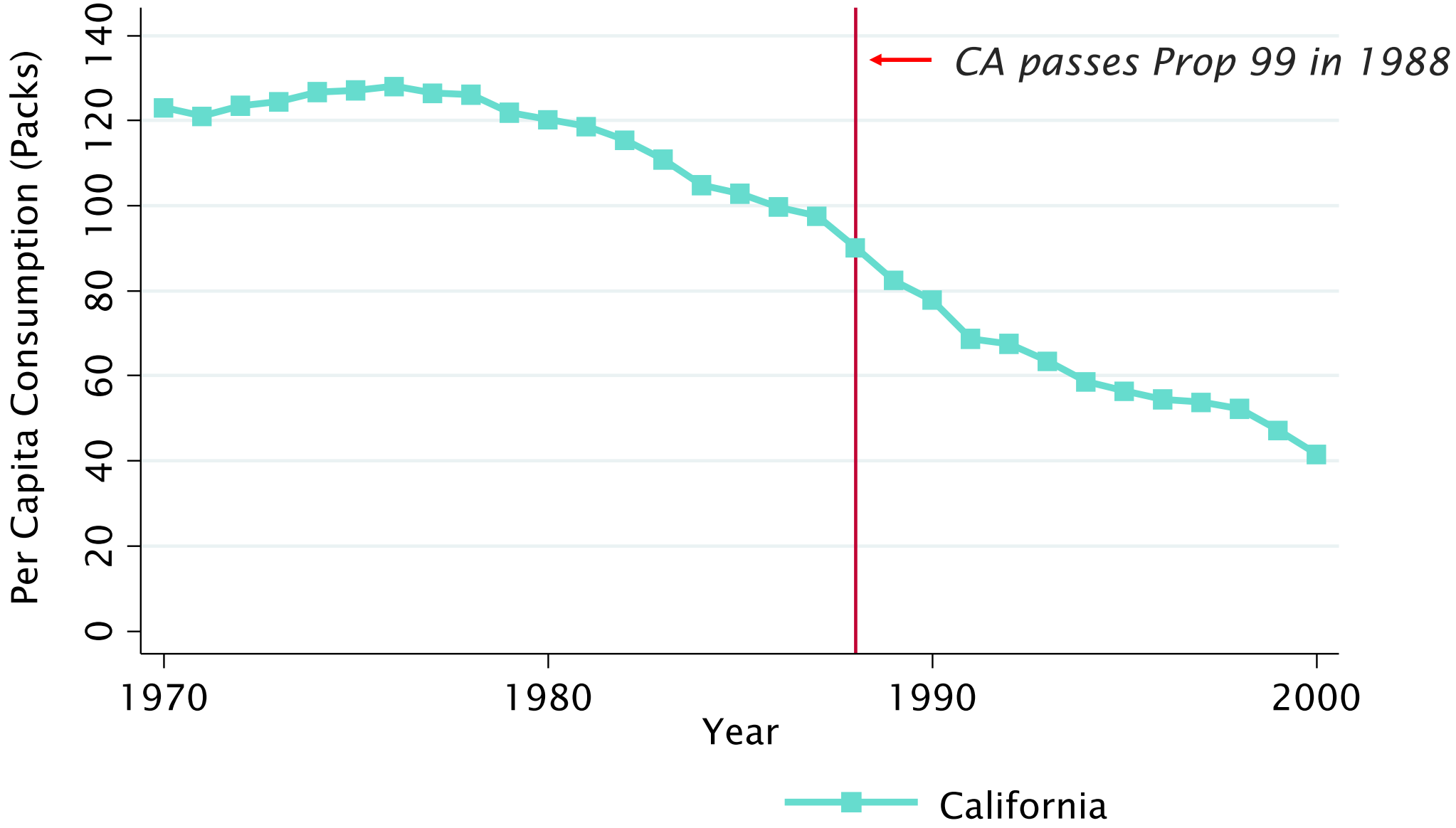


Policies to Improve Life Expectancy: California Proposition 99

- In 1988, California passed the first large-scale tobacco control program in the United States as Proposition 99:
  - Increased California's cigarette excise tax by 25 cents per pack
  - Earmarked the tax revenues to health and anti-smoking education budgets
  - Funded anti-smoking media campaigns
  - Spurred local clean indoor-air ordinances throughout the state
- What was the effect of Proposition 99 on cigarette consumption in CA?
- We will start with a standard differences in differences approach

What was the effect of California's Proposition 99 on smoking?

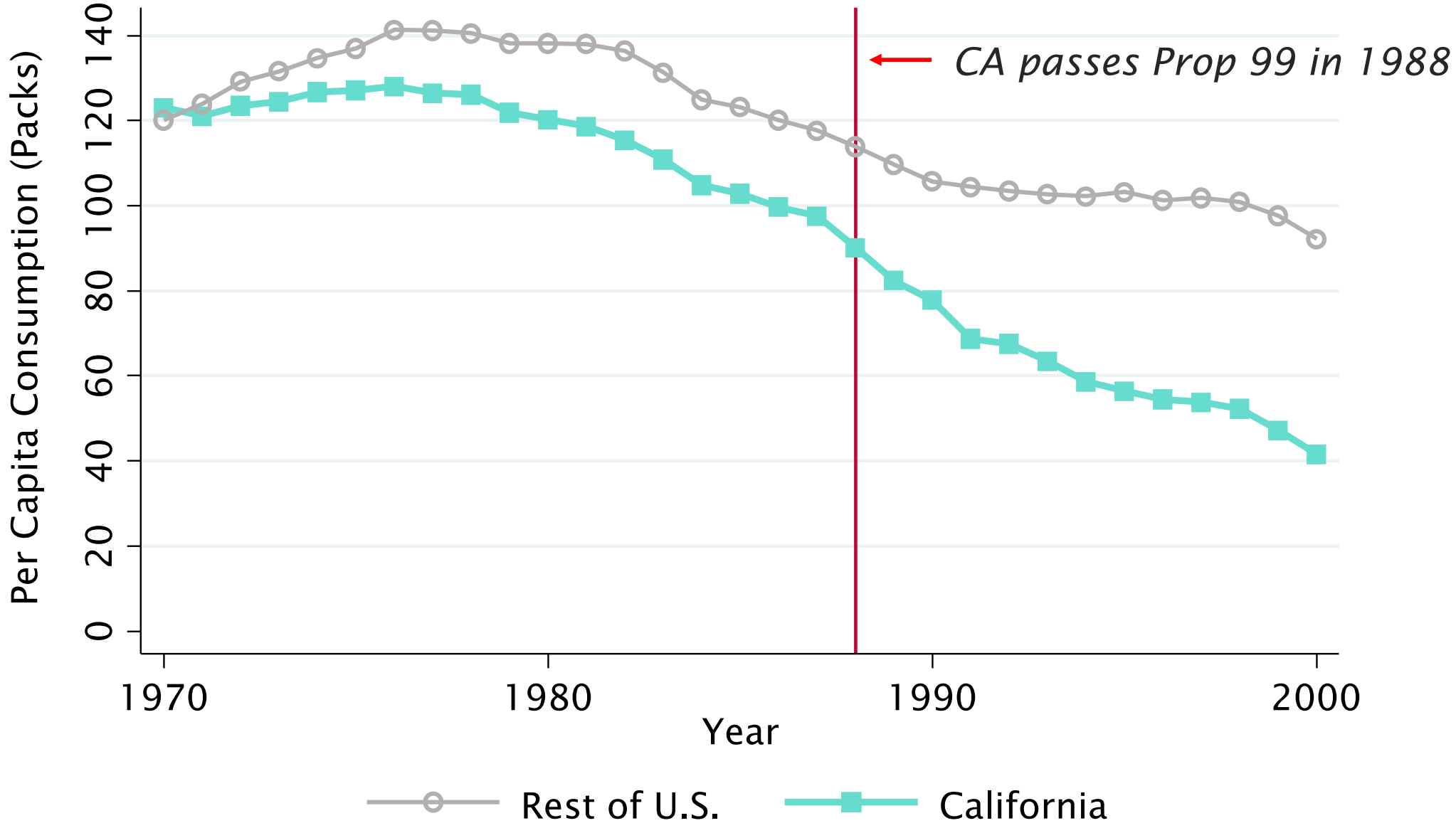
Cigarette Consumption in California and rest of the U.S.



Source: Bruich 2020 replication of Abadie, Diamond, and Hainmueller (2010) figure 1

What was the effect of California's Proposition 99 on smoking?

Cigarette Consumption in California and rest of the U.S.



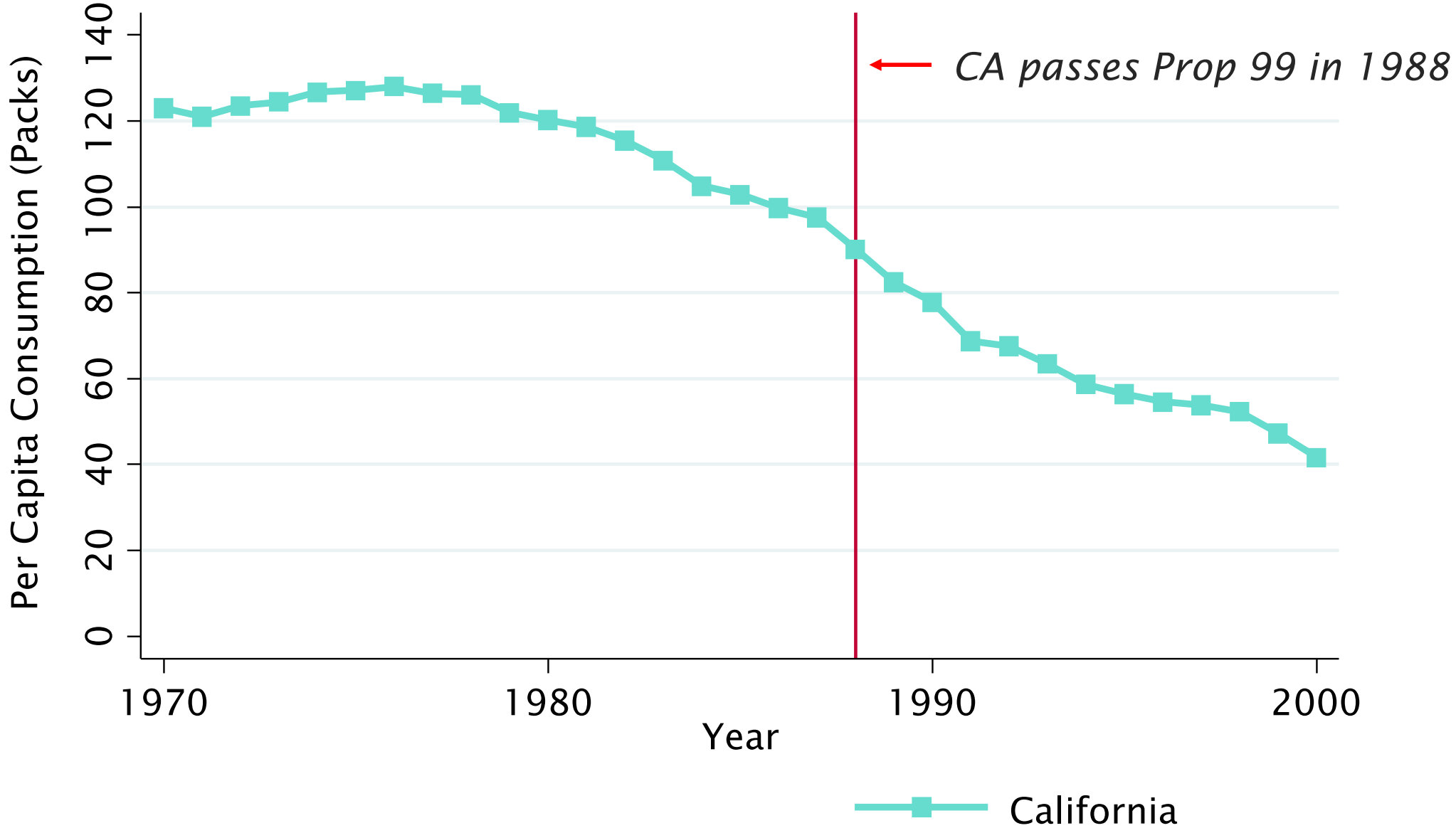
Source: Bruich 2020 replication of Abadie, Diamond, and Hainmueller (2010) figure 1

Synthetic Control Method

- The main difficulty for differences in differences is identifying an appropriate control group that satisfies the parallel trends assumption
- The **synthetic control method** is the state-of-the-art approach for choosing a control group for "comparative case studies"
- Instead of comparing California to all other states or nearby states, we construct a "synthetic California" that is a weighted average of other states
- Weights are chosen using data-driven procedure with many set to 0
- Synthetic Control Method is most useful with aggregate data
- Often used to complement standard differences in differences methods

What was the effect of California's Proposition 99 on smoking?

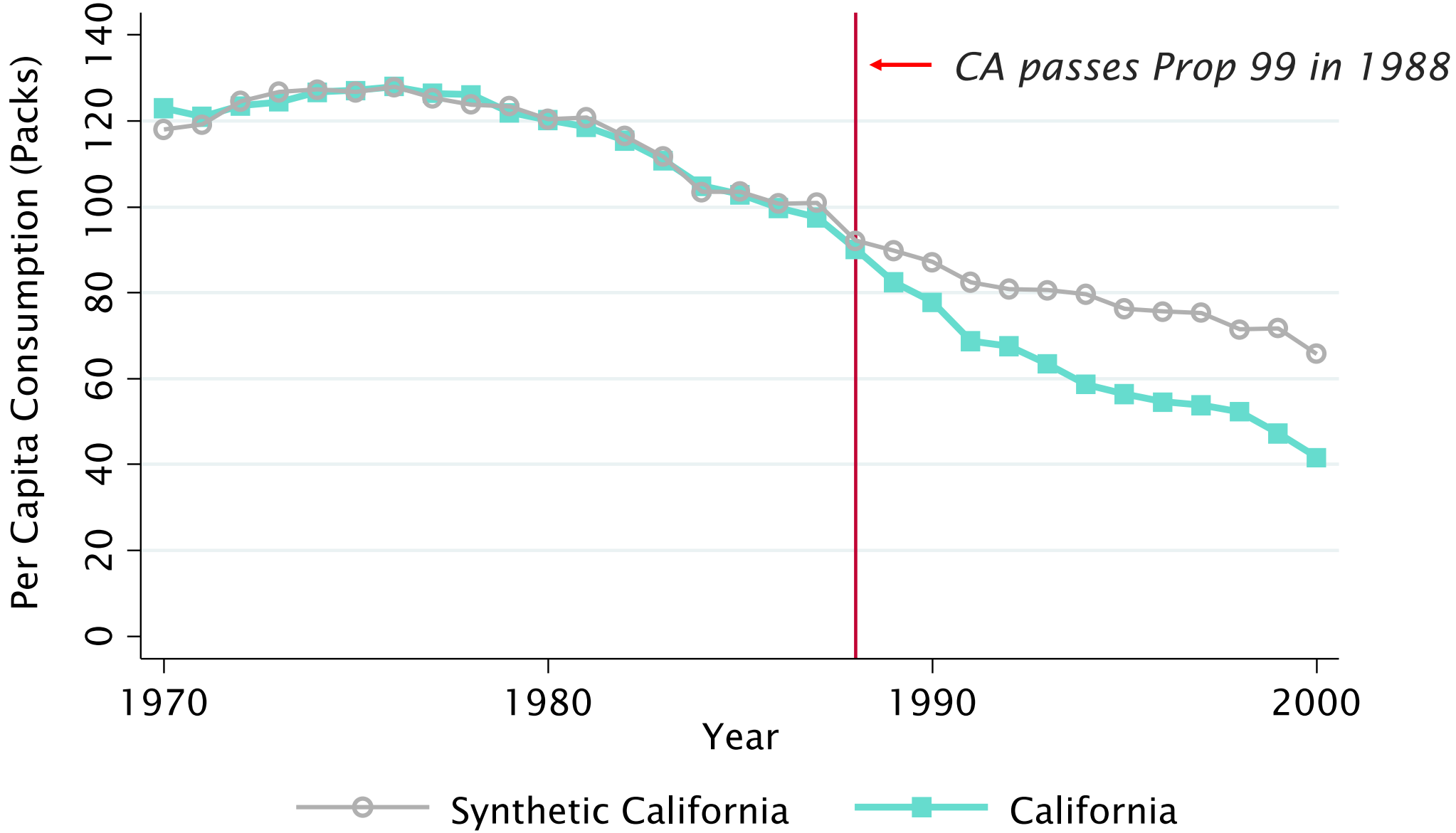
Cigarette Consumption in California and Synthetic California



Source: Bruich 2020 replication of Abadie, Diamond, and Hainmueller (2010) figure 2

What was the effect of California's Proposition 99 on smoking?

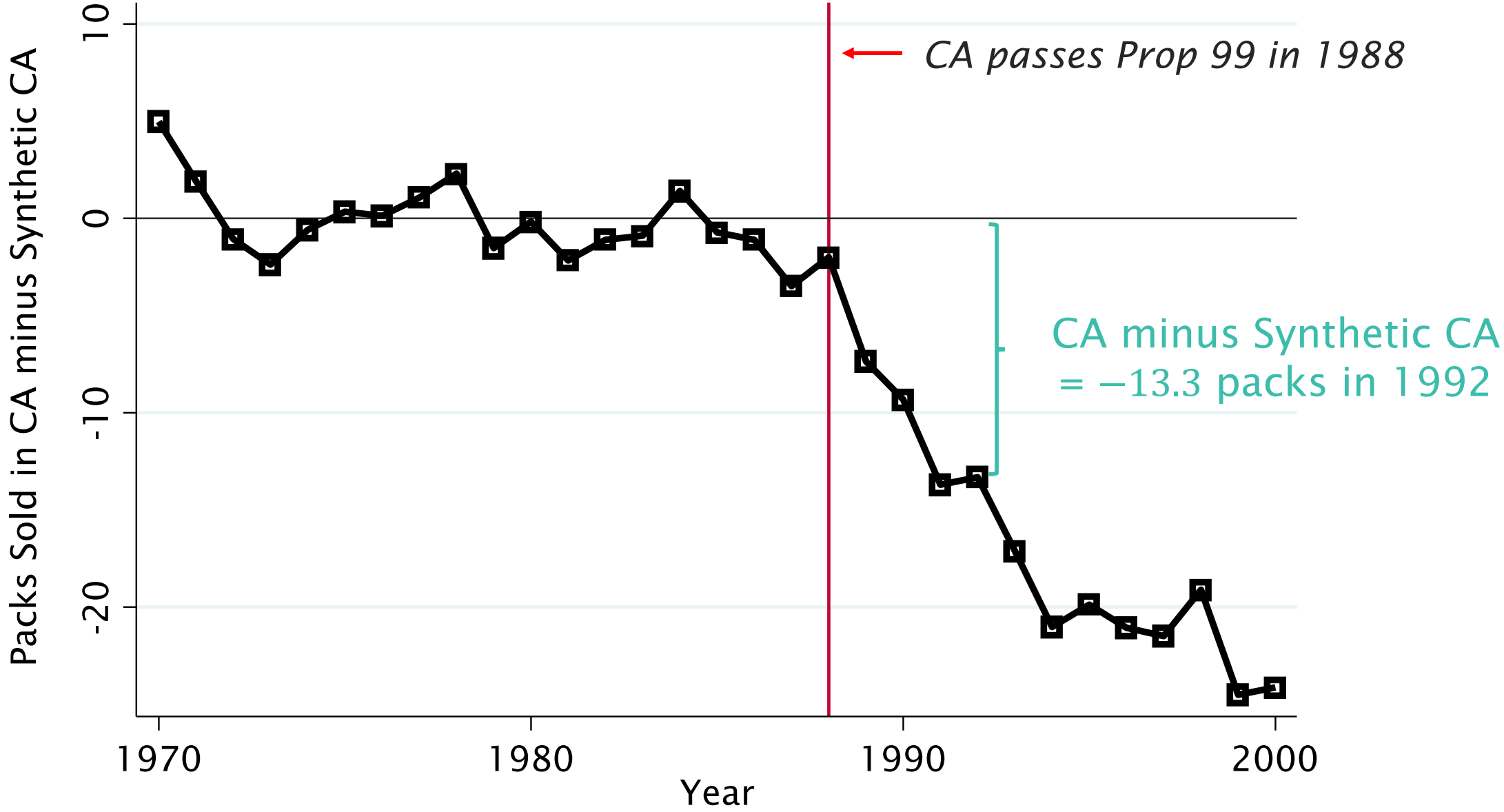
Cigarette Consumption in California and synthetic California



Source: Bruich 2020 replication of Abadie, Diamond, and Hainmueller (2010) figure 2

What was the effect of California's Proposition 99 on smoking?

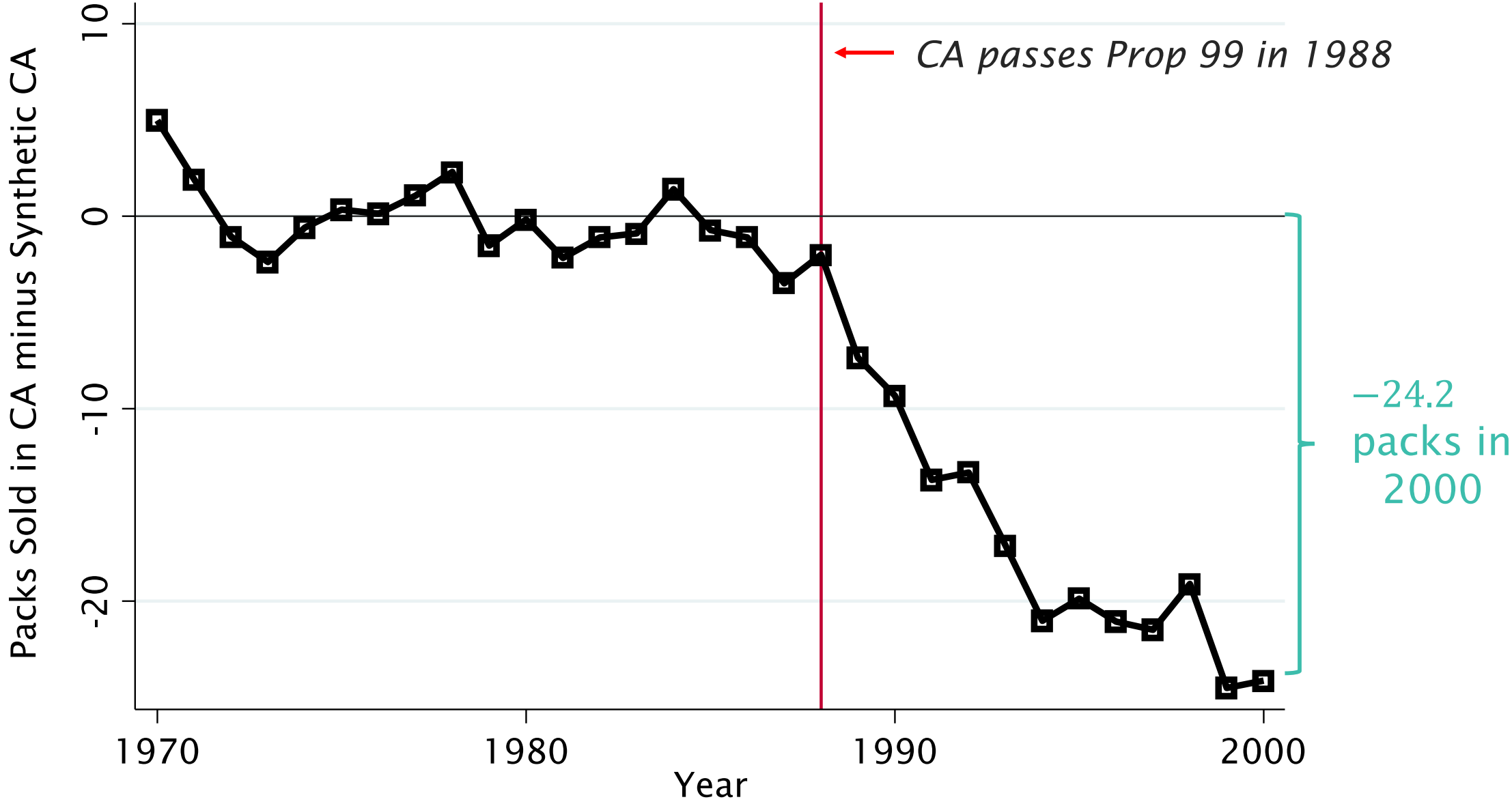
Cigarette Consumption in California minus synthetic California



Source: Bruich 2020 replication of Abadie, Diamond, and Hainmueller (2010) figure 3

What was the effect of California's Proposition 99 on smoking?

Cigarette Consumption in California minus synthetic California



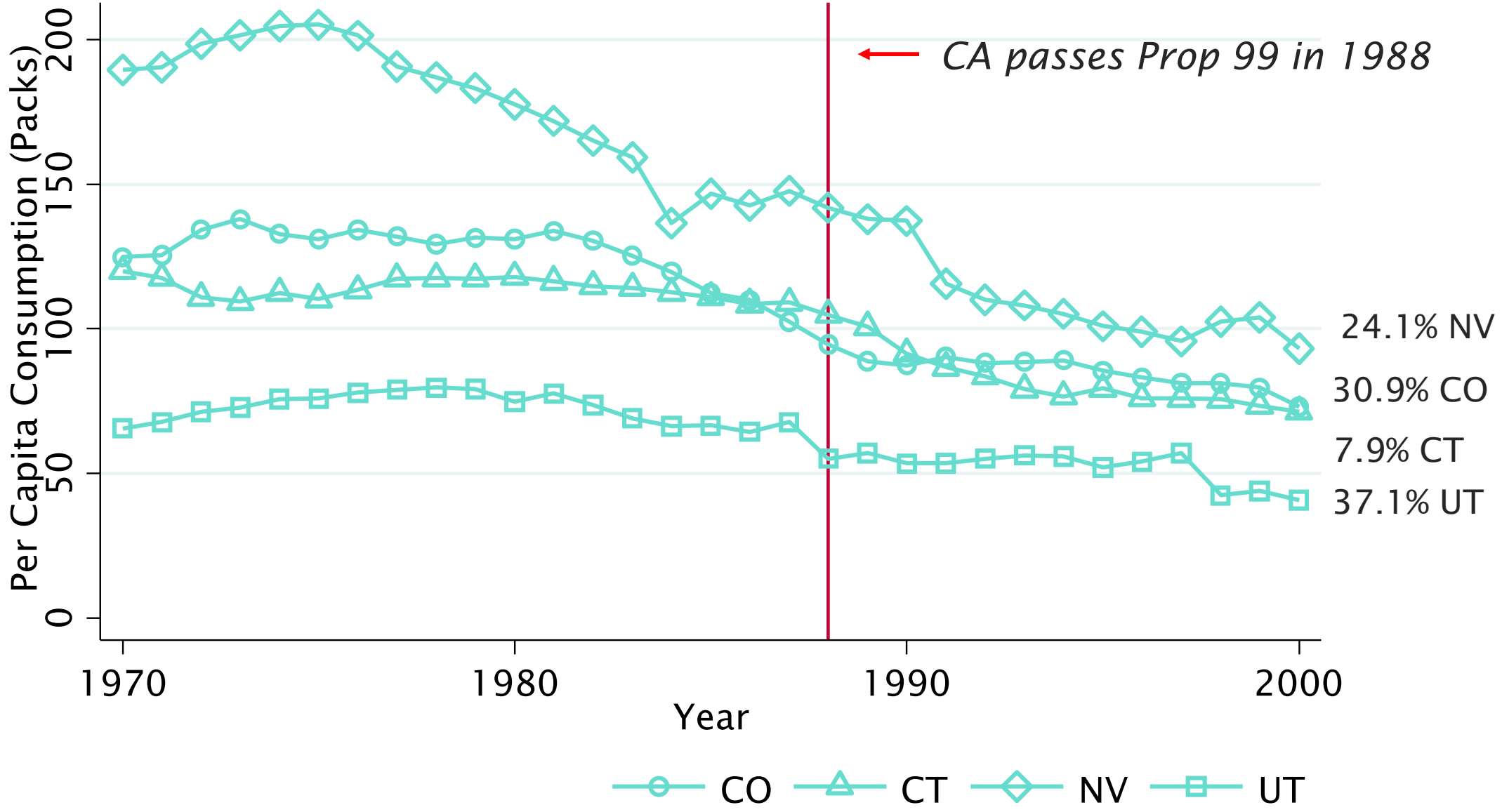
Source: Bruich 2020 replication of Abadie, Diamond, and Hainmueller (2010) figure 3

What is Synthetic California?

- Synthetic California is a weighted average of per capita cigarette consumption in these other states:
  - 37.1% cigarette consumption in Utah
  - 30.9% cigarette consumption in Colorado
  - 24.1% cigarette consumption in Nevada
  - 7.9% cigarette consumption Connecticut
- All other states are given 0% weight
- Although none of these states are comparable with CA by themselves, the weighted average lines up closely

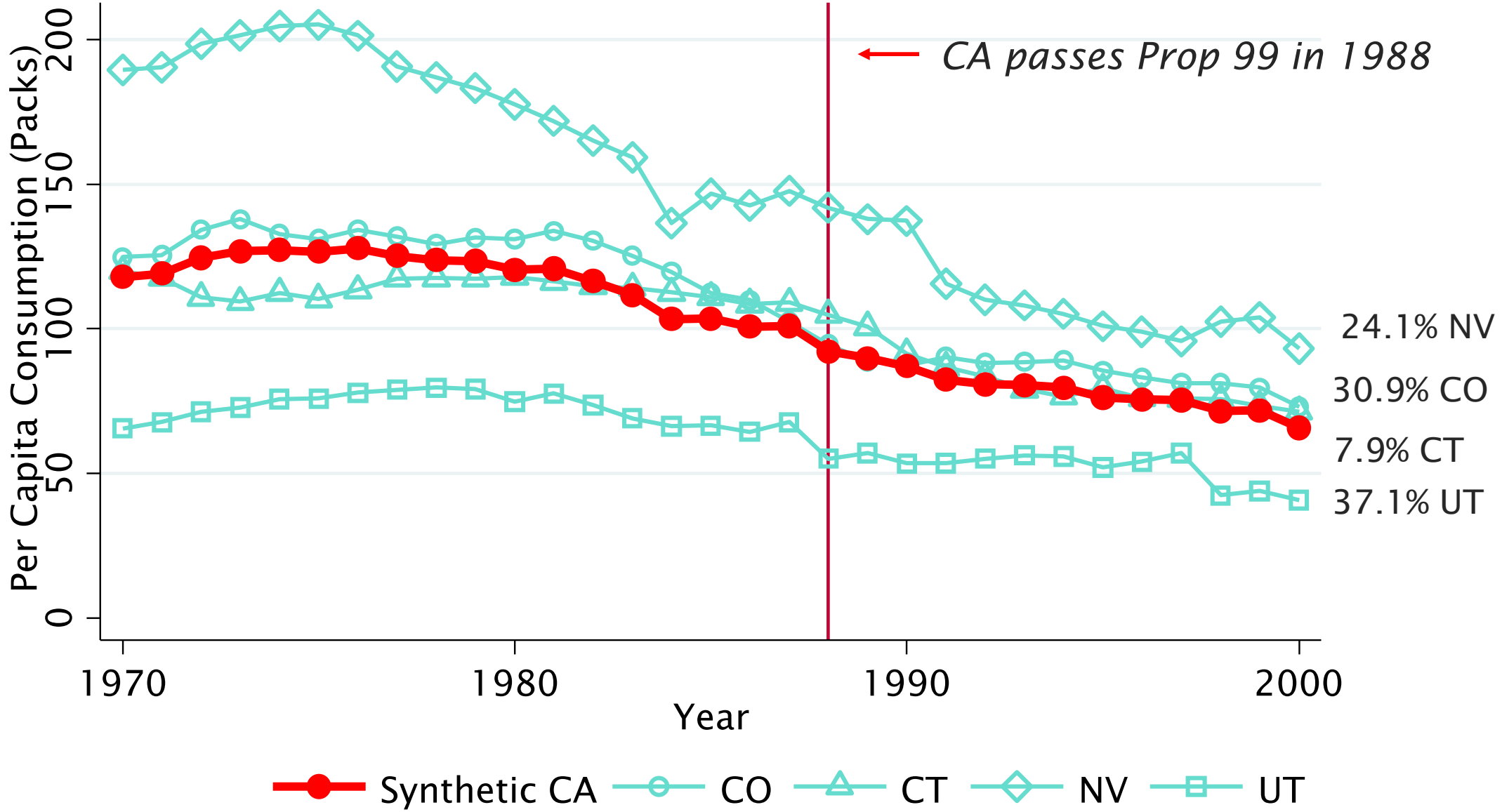
What is Synthetic California?

Cigarette Consumption in Utah, Colorado, Nevada, and Connecticut



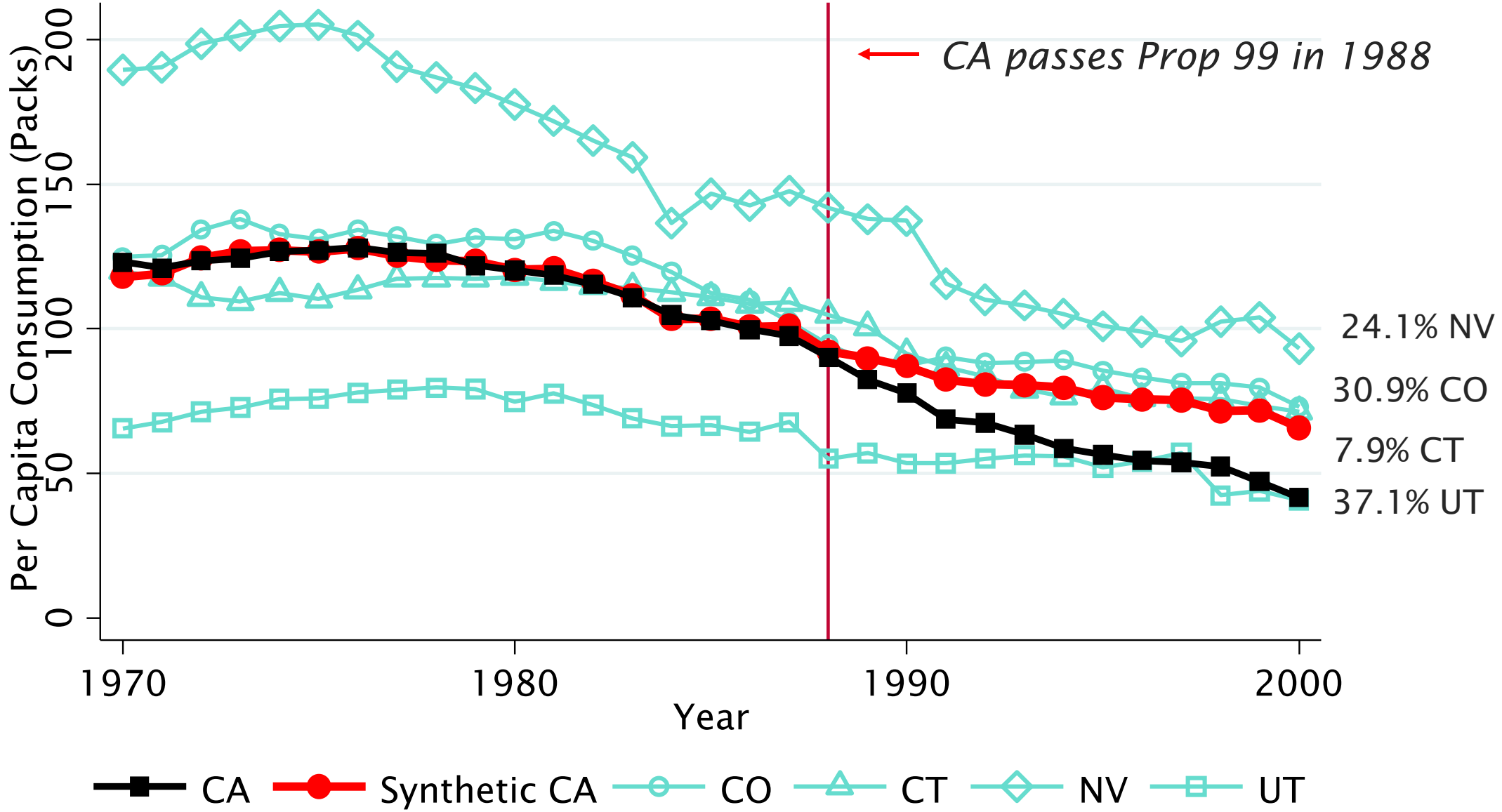
What is Synthetic California?

Cigarette Consumption in Utah, Colorado, Nevada, and Connecticut



What is Synthetic California?

Cigarette Consumption in Utah, Colorado, Nevada, and Connecticut



Weights in the Synthetic Control Method

- Weights are non-negative and sum to 1
- The weights are chosen to minimize the “distance” between CA and Synthetic CA with respect to pre-treatment period characteristics
- For synthetic control, the *most important* pre-treatment period characteristics are the lags of the dependent (outcome) variable
- The lags of the dependent variable are cigarette consumption in 1970-1987
- In practice, the number of predictor variables has to be kept small

Balance of predictor variables used to construct weights

Variable	California	Synthetic California	Rest of U.S.
Gallons of Beer Consumed	23.7	23.3	23.9
log(Personal Income)	9.3	9.2	9.1
Cost per pack of cigarettes	$0.71	$0.70	$0.69
Age 15-24	18.1%	18.4%	18.1%

Balance of predictor variables used to construct weights

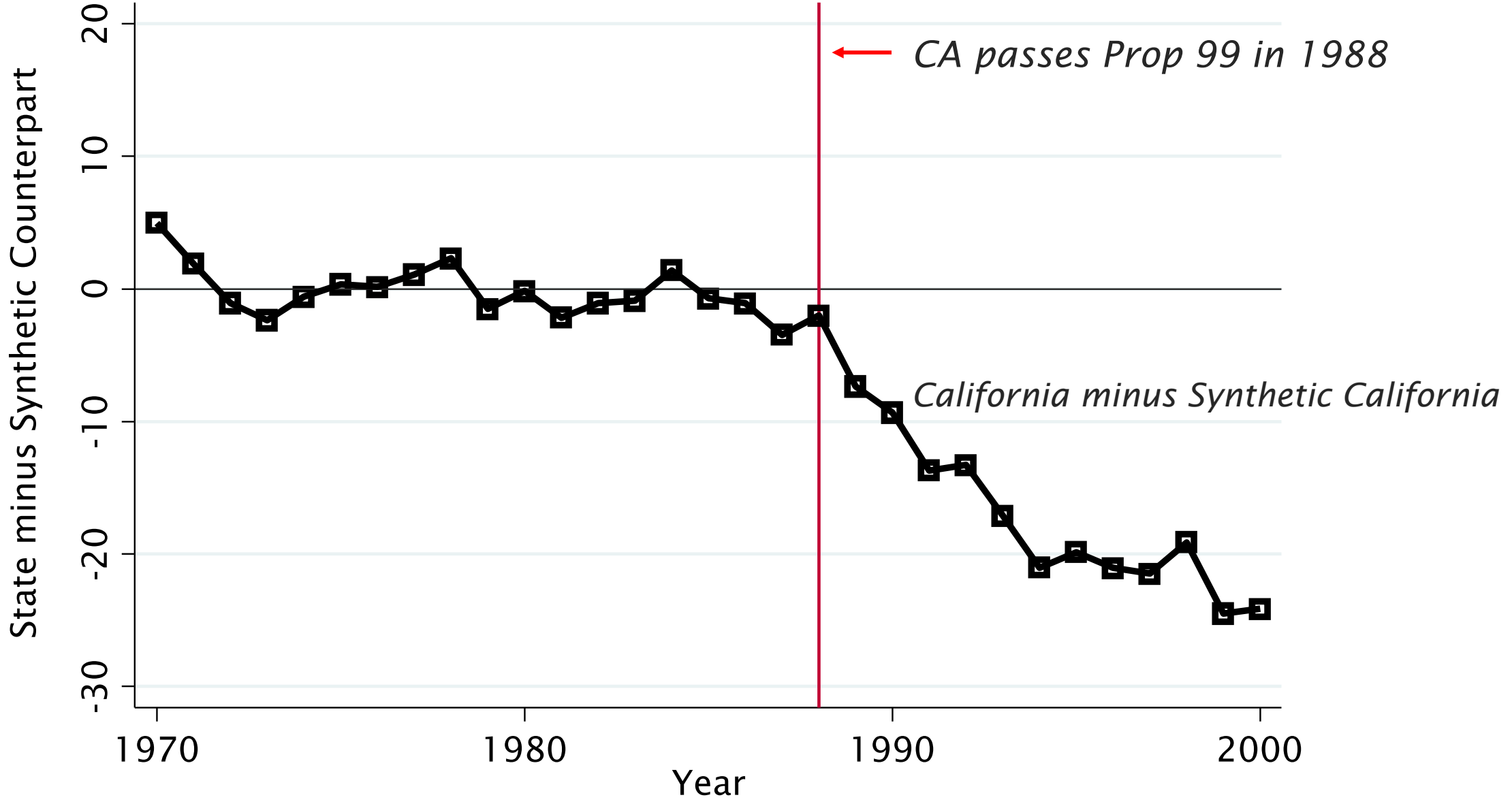
Variable	California	Synthetic California	Rest of U.S.
Gallons of Beer Consumed	23.7	23.3	23.9
log(Personal Income)	9.3	9.2	9.1
Cost per pack of cigarettes	$0.71	$0.70	$0.69
Age 15-24	18.1%	18.4%	18.1%
Cigarette consumption in 1987	90.1	92.1	113.8
Cigarette consumption in 1980	120.2	120.4	138.1
Cigarette consumption in 1975	127.1	126.8	136.9

Permutation tests

- To quantify uncertainty around the choice of Synthetic Control, we use **permutation tests**
- Intuitively, the source of uncertainty comes from the **unknown counterfactual**
- In a permutation test, we reassign the treatment to each of the units in the donor pool
  - **Donor pool**: units that are potential controls for constructing the synthetic control
- We estimate a "placebo treatment effect" for each unit in the donor pool
- The effect of the treatment on the unit affected by the intervention is deemed to be significant when its magnitude is extreme relative to the distribution of the "placebo treatment effects"

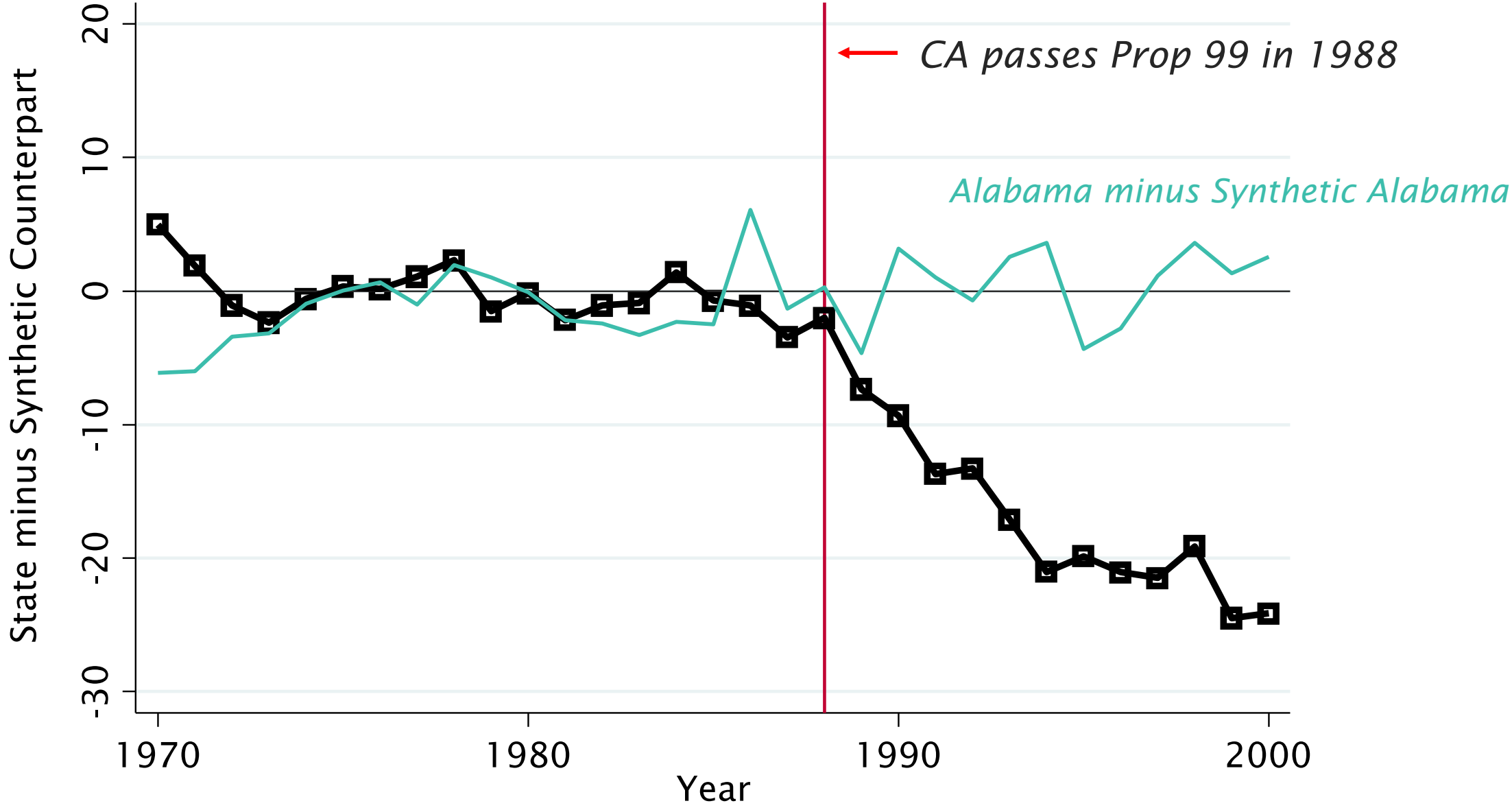
Permutation test: "Placebo in Space"

Cigarette consumption in each state minus its synthetic counterpart



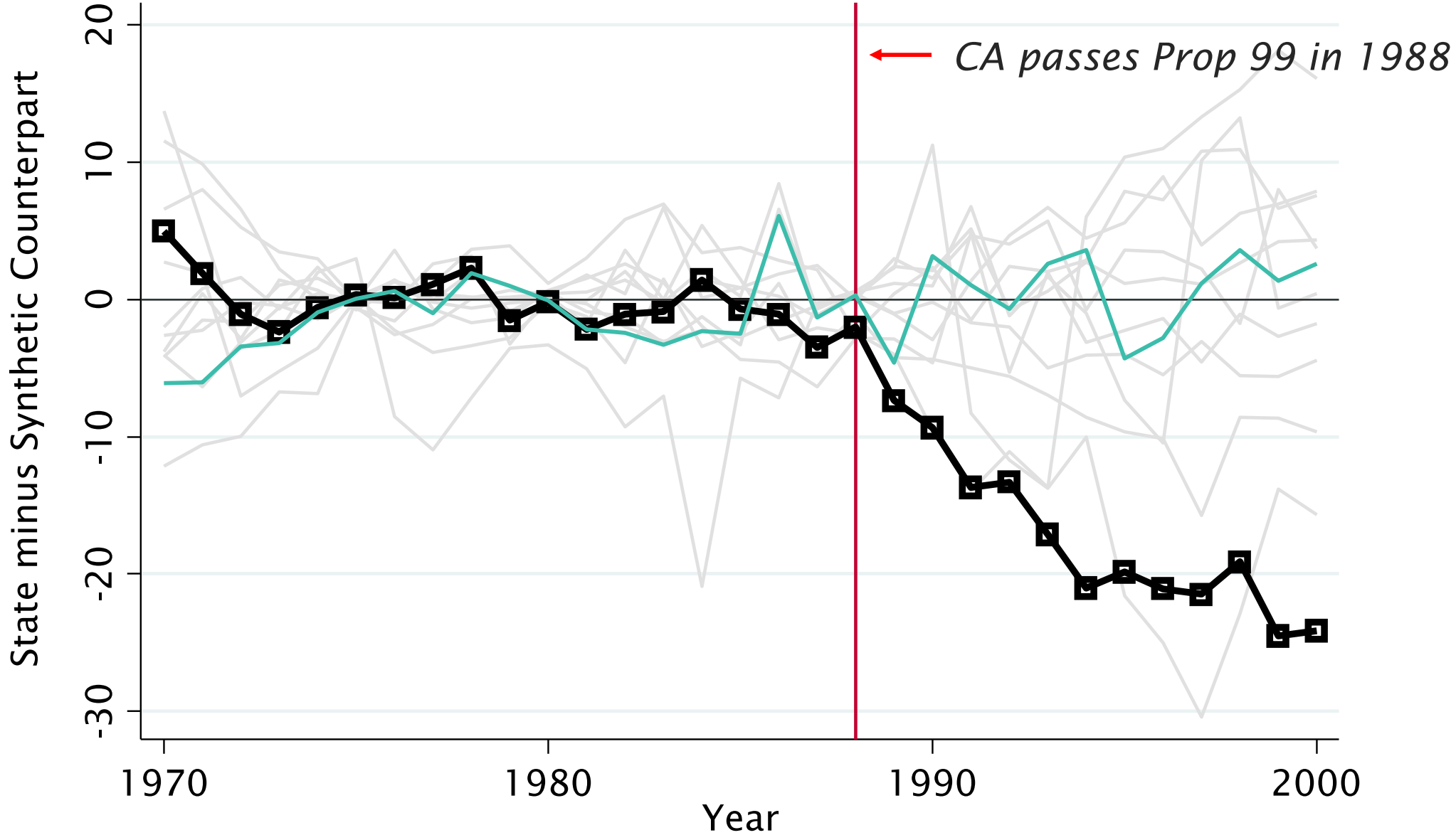
Permutation test: "Placebo in Space"

Cigarette consumption in each state minus its synthetic counterpart



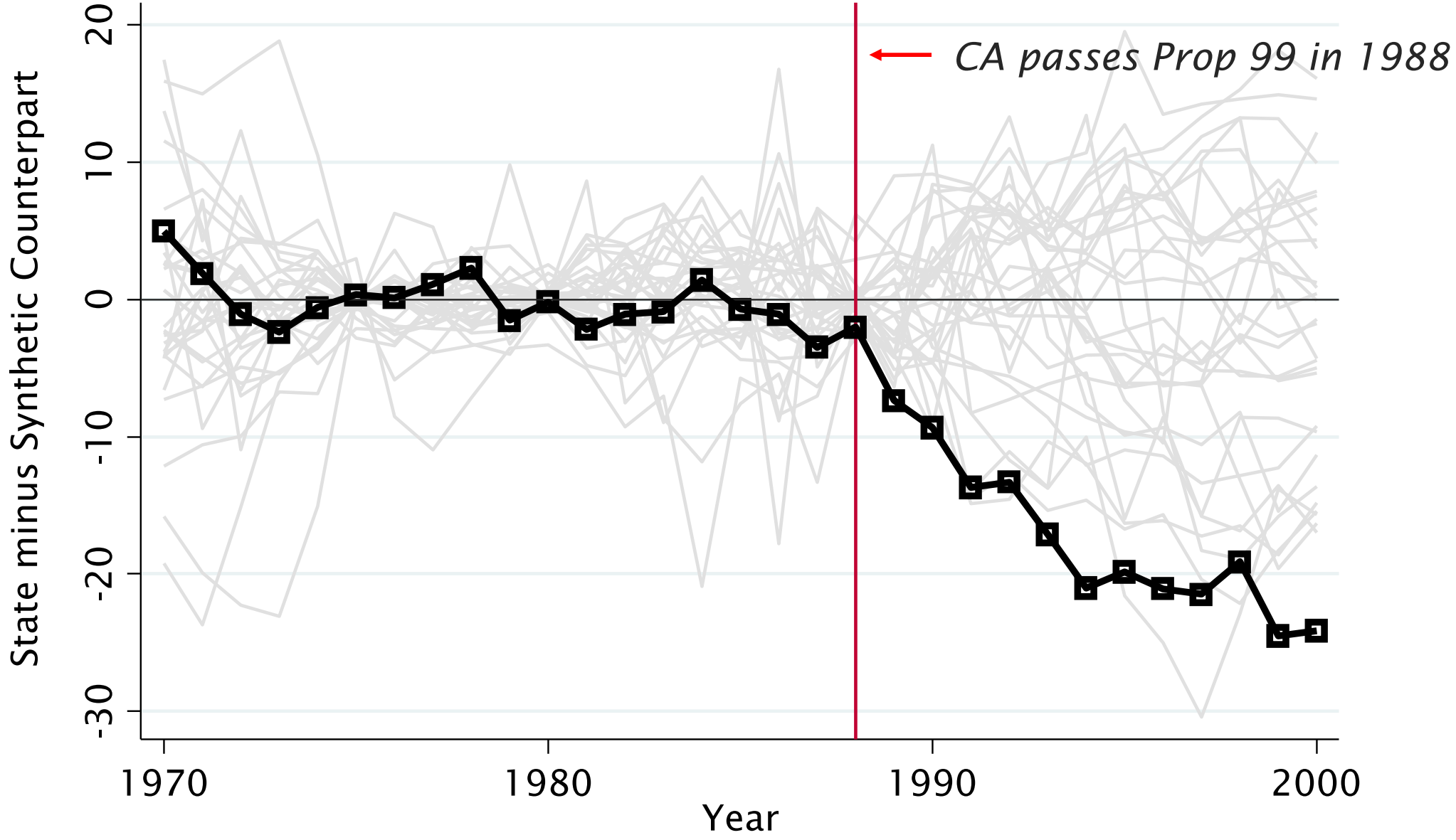
Permutation test: “Placebo in Space”

Cigarette consumption in each state minus its synthetic counterpart



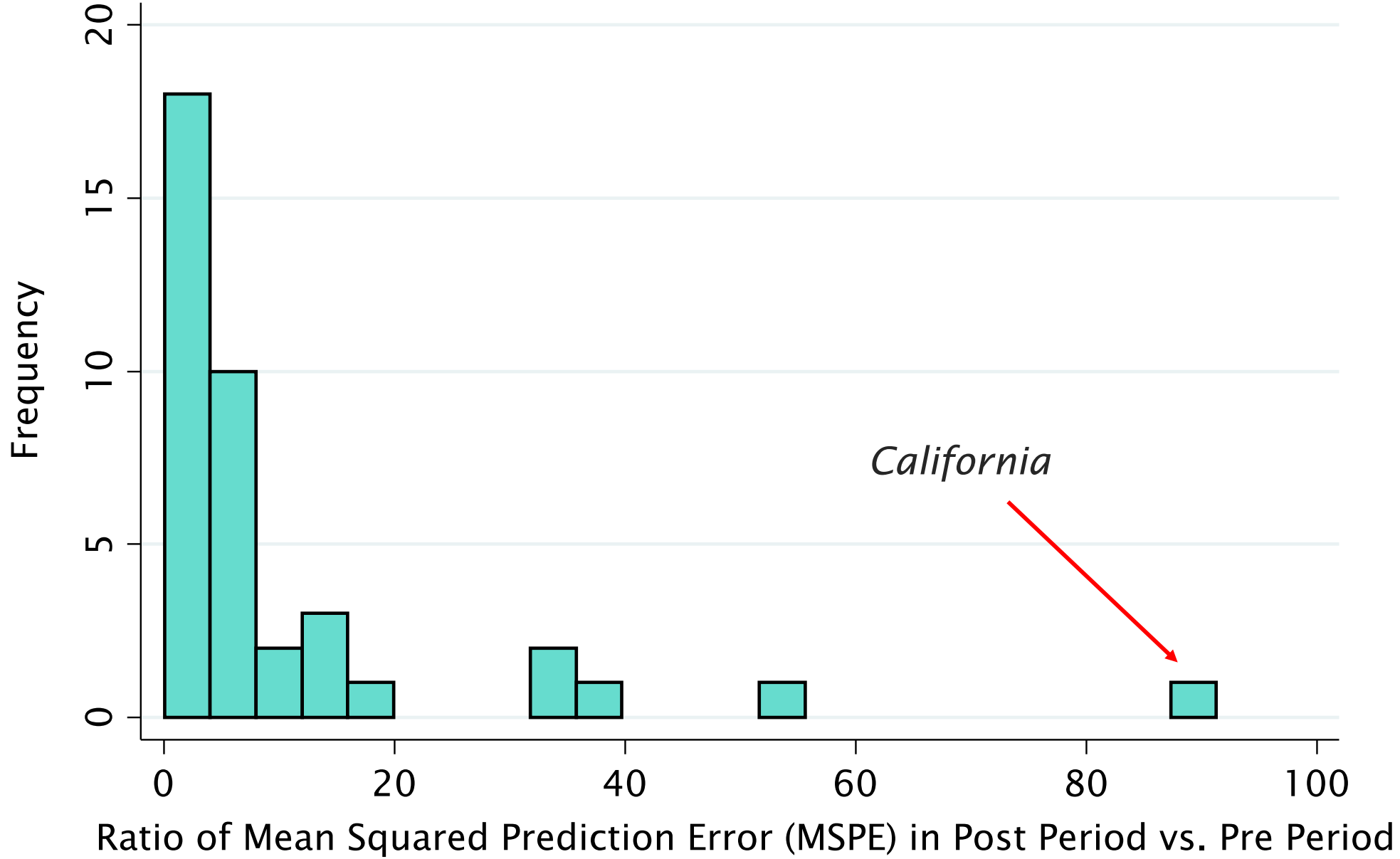
Permutation test: "Placebo in Space"

Cigarette consumption in each state minus its synthetic counterpart



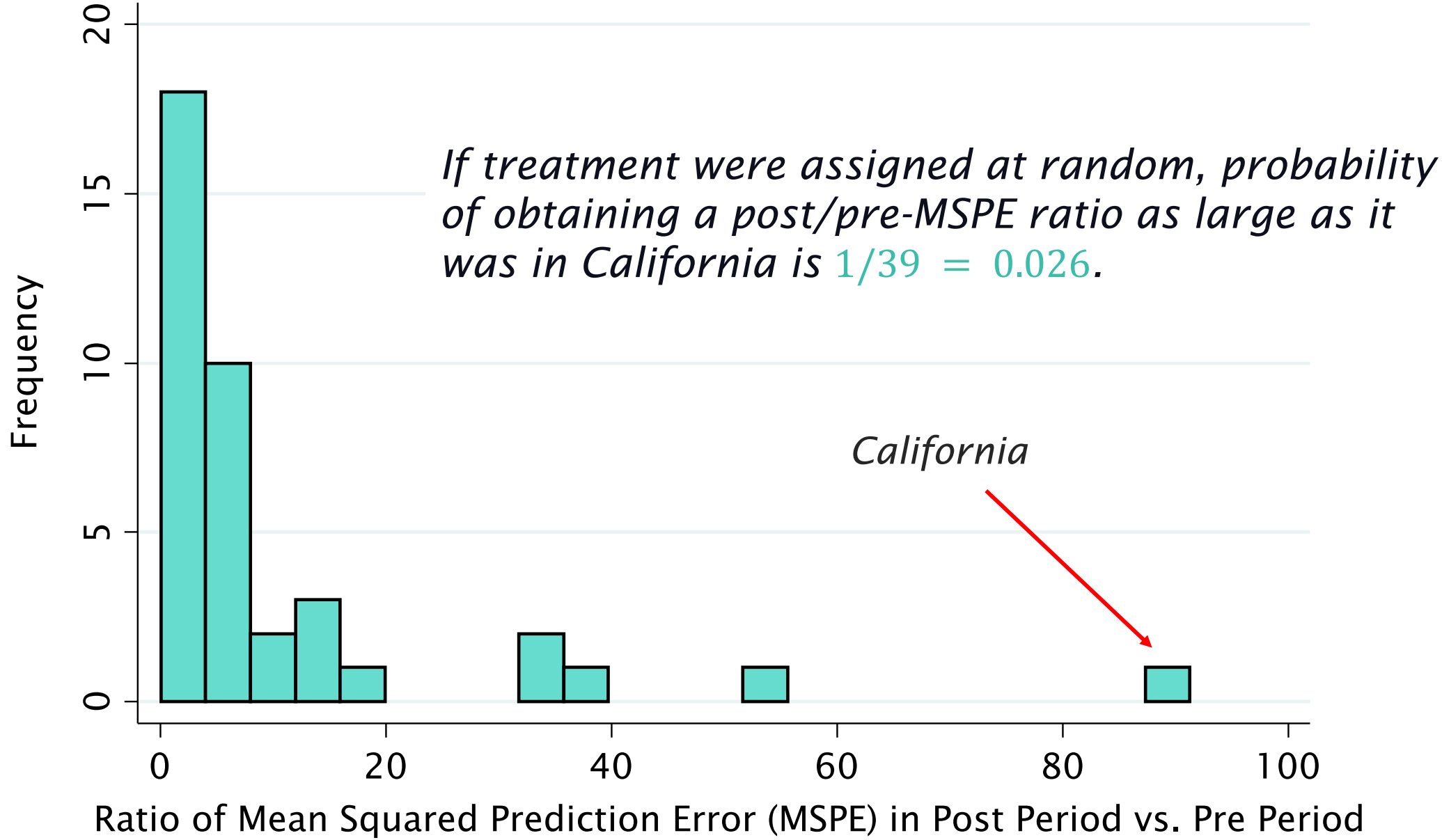
Distribution of Placebo Estimates from Permutation Test

Ratio of Post Prop 99 MSPE and Pre Prop 99 MSPE: California and 38 Control States



Distribution of Placebo Estimates from Permutation Test

Ratio of Post Prop 99 MSPE and Pre Prop 99 MSPE: California and 38 Control States



Data requirements

- The treatment should be at an aggregate level, such as a city or a state
  - Most useful with aggregate data on predictors and outcomes, as opposed to estimates constructed from survey data
- Synthetic control method is not designed for individual level treatment
- The credibility of synthetic control method depends on its ability to steadily track the trajectory of the outcome variable for the affected unit before the intervention
- Therefore, data should be available for a *long* period of time before the intervention or treatment occurred

The Washington Post

Wonkblog

Seriously, here's one amazing math trick to learn what can't be known

THE WALL STREET JOURNAL

REAL TIME ECONOMICS | ECONOMICS

How an Analysis of Basque Terrorism Helps Economists Understand Brexit

A method pioneered by an MIT professor has also been used to estimate the economic effect of a tobacco ban, German reunification, legalization of prostitution and gun rights

Source: Abadie (2020)

Take aways

- **Synthetic Control Method** is the state-of-the-art approach to select a comparison group for differences in differences
- The idea is to avoid cherry-picking in selecting the control group
- We use data-driven procedure to construct a weighted average of several control units, with many weights set equal to zero
- Both synthetic control and differences in differences methods rely on the same parallel trends identification assumption
- *"The synthetic control approach… is arguably the most important innovation in the policy evaluation literature in the last 15 years"* (Athey and Imbens 2017)