Synthetic control

DENG KUANGKUANG

What's different about panel data?

- The fundamental problem of causal inference
- A statistical solution makes use of the population

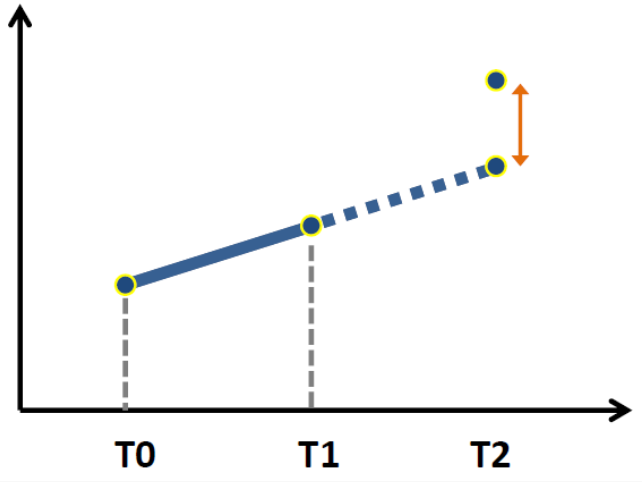
$$\text{e.g. } T = E[Y_1] - E[Y_0]$$

- A scientific solution exploits homogeneity or invariance assumptions
  - search for patterns across different cases
  - When something happens, some particular outcome will follow
  - E.g. The long-run growth rate of the US economy is 2.5%, (while the growth rate in each year fluctuates).

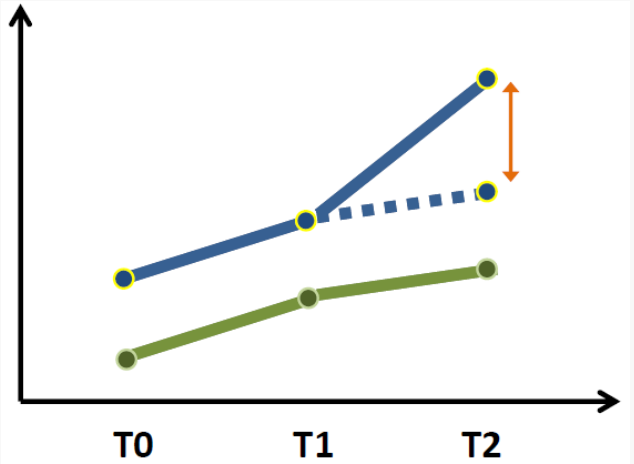
- Panel data allow us to construct the counterfacturals of the treated units in the post-treatment period using information from both the control group and the treatment group in the pre-treatment period.

Causal inference with panel data

Time-series analysis



Causal inference with panel data



- Given aggregate data, an external shock happened at T1, what is the outcome?

- Statistical solution: matching
- Scientific solution: modeling
- Panel data make both easier
  - Matching on lagged outcomes makes matching more plausible
  - Parallel trends assumption is somewhat “testable”

Theoretical motivation

- Fixed effects (or DID) model

$$Y_{1t}^{N} = X_{it}\beta + Z_i\theta_t + \delta_t + \alpha_i + \epsilon_{it}$$

  - Consider the outcome of a property tax on housing price in Shanghai
  - DID: compare housing price in Beijing and Shanghai in pre- and post- tax periods
  - Assumption: without the treatment, the trends of HP in BJ and SH should be the same (comparability). But this may not be the case! E.g. the effects of interest rate on HP may vary with land supply, and land supply plans differ in BJ and SH. => interest rate ↓ 1%, $HP_{beijing}$ ↑10% v.s. $HP_{shanghai}$ ↑5%

- What if the true model is as complicated as:

$$Y_{1t}^{N} = X_{it}\beta + Z_i\theta_t + \lambda_t\mu_i + \epsilon_{it}$$

- $\lambda_t\mu_i$ are fixed effects interacted with time-varying coefficient, in which δ_t and α_i are special cases

- Abadie and Gardeazabal (2003), Abadie, Diamond, and Hainmueller (2010) found a solution when there is only one treated unit

Comparative case studies

Comparative case studies:

- Compare the evolution of an aggregate outcome for the unit affected by the intervention to the evolution of the same aggregate for some control group (e.g. Card, 1990, Card and Krueger, 1994, Abadie and Gardeazabal, 2003)
- Events or interventions take place at an aggregate level (e.g. cities, provinces, countries).

Challenges:

- N_{tr} is small by definition
- Selection of control group is often ambiguous
- Standard errors do not reflect uncertainty about the ability of the control group to reproduce the counterfactural of interest

The synthetic control method: setup

- Suppose that we observe J+1 regions in periods 1, 2, …, T.
- Let T_0 be the number of pre-intervention periods. Region "one" is exposed to the intervention during periods T_0+1, …, t.
- Let Y_{it}^N be the outcome that would be observed for region i at time t in the absence of the intervention.
- Let Y_{it}^I be the outcome that would be observed for region i at time t if region i is exposed to the intervention in periods T_0+1 to T.
- We aim to estimate the effect of the intervention on the treated unit $(\alpha_{1T_0+1}, \dots, \alpha_{1T})$, where $\alpha_{1t} = Y_{1t}^I - Y_{1t}^N = Y_{1t} - Y_{1t}^N$ for $t > T_0$

observed

Needs to be estimated: the outcome of the treated regions if it were not treated

Setup

- Suppose Y_{1t}^N is given by a factor model:

$$Y_{1t}^N = \delta_t + \boldsymbol{\theta_t Z_i} + \lambda_t \mu_i + \epsilon_{it} \quad (1)$$

  - δ_t is an unknown common time-dependent factor with constant factor loadings across units
  - $\boldsymbol{Z_i}$ is a (1 × r) vector of observed covariates,
  - $\boldsymbol{\theta_t}$ is a (r × 1) vector of unknown parameters,
  - λ_t is a (1 × F) vector of unknown common factors,
  - μ_i is a (F × 1) vector of unknown factor loadings.

- $\lambda_t \mu_i$: heterogeneous responses to multiple unobserved factors

- Basic idea: reweight the control group such that the synthetic control unit matches $\boldsymbol{Z_i}$ and (some) pre-treatment Y_{1t} of the treated units; then μ_i is automatically matched.

Theory

- Let $W = (w_2, \dots, w_{J+1})'$ with $w_j \geq 0$ for j = 2, … , J+1 and $w_2, + \cdots, + w_{J+1} = 1$. Each value of W represents a potential synthetic control, i.e. a particular weighted average of control regions

- Suppose there are $W^* = (w_2^*, \dots, w_{J+1}^*)$ such that

$$\sum_{j=2}^{J+1} w_j^* Y_{j1} = Y_{11}, \quad \sum_{j=2}^{J+1} w_j^* Y_{j2} = Y_{12}, \dots, \quad .$$
$$\sum_{j=2}^{J+1} w_j^* Y_{jT_0} = Y_{1T_0}, \quad \text{and } \sum_{j=2}^{J+1} w_j^* Z_j = Z_1 \quad (2)$$

- For each of the T_0 pre-intervention values of the outcome of the treated, W^* finds a linear combination of the outcomes of the untreated that equals the outcome of the treated. W^* does the same for the time invariant $\boldsymbol{Z_i}$.

Illustrative example

- The effects of property tax on HP in Shanghai
- The policy implemented in Jan 2011
- Select a group of control regions, e.g. the other 69 big cities
- Suppose the study period is Jan 2010 to Jan 2012
- Suppose HP is also affected by the size of population (can be regarded as time-invariant within a short time)

- The theory:
  - Find a matrix of weight $W^* = \left(w_2^*, \dots, w_{J+1}^*\right)$
  - Such that the linear combination of the HP in the 69 control cities equal to the HP in Shanghai for Jan 2010, Feb 2020, … Dec 2010
  - And the same linear combination of population in the 69 control cities equal to that in Shanghai

The theory

- W^* can only be found exactly only if $(Y_{11}, \dots Y_{1T_0}, Z_1)$ belongs to the convex hull of $\{(Y_{21}, \dots Y_{2T_0}, Z_2), \dots, (Y_{J+11}, \dots Y_{J+1T_0}, Z_{J+1})\}$
  - The housing price in Shanghai was almost the highest within the 70 cities
  - The population in Shanghai has been the largest within the 70 cites
  - => convex hull requirement unsatisfied.

- In practice, it is often the case that no set of weights exists such that Equation (2) holds. e.g. in the case of Shanghai HP.
- In some instances, the fit may be poor and then a synthetic control is not recommended.

- But if W^* exists, an approximately unbiased estimator of α_{1t} is

$$\widehat{\alpha_{1t}} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

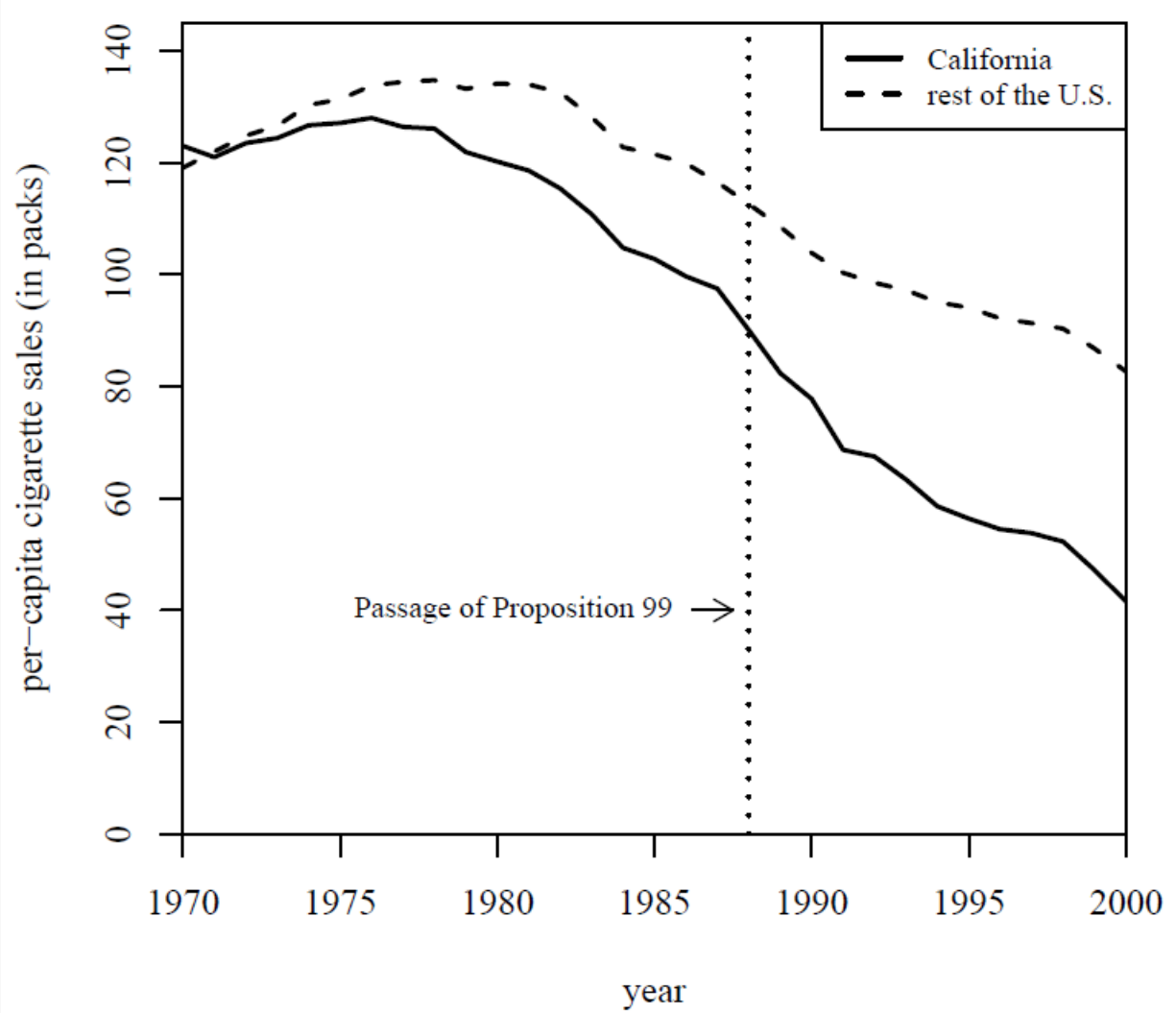
Application 1: California's Proposition 99

In 1988, California first passed comprehensive tobacco control legislation:

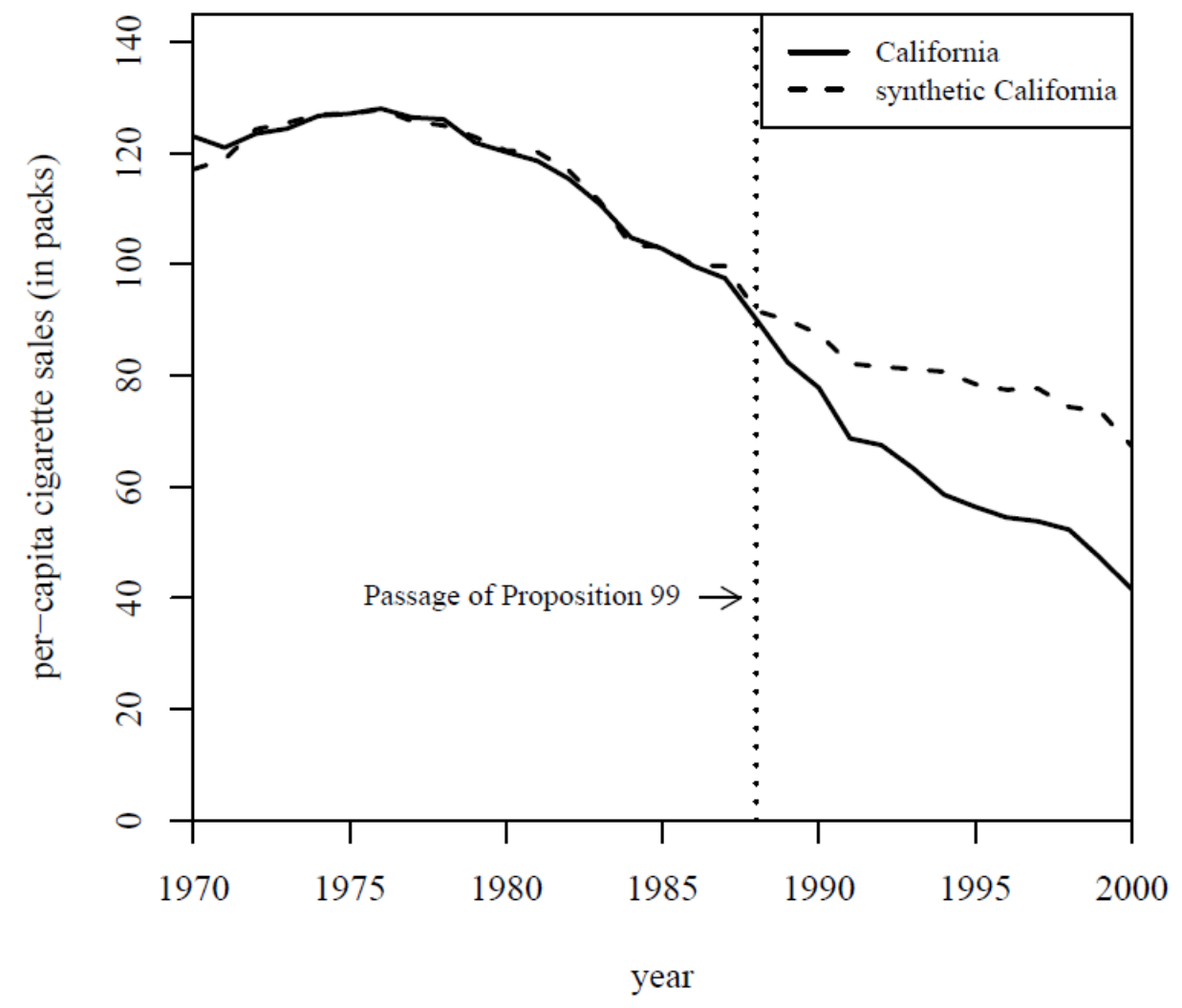
- Increased cigarette tax by 25 cents/pack
- Earmarked tax revenues to health and anti-smoking budgets
- Funded anti-smoking media campaigns
- Spurred clean-air ordinances throughout the state
- Produced more than $100 million per year in anti-tobacco projects

Other states that subsequently passed control programs are excluded from the donor pool of controls

Cigarette consumption: CA and the rest of the U.S.



Cigarette consumption: CA and synthetic CA



Weights of the controls

Table 2. State weights in the synthetic California

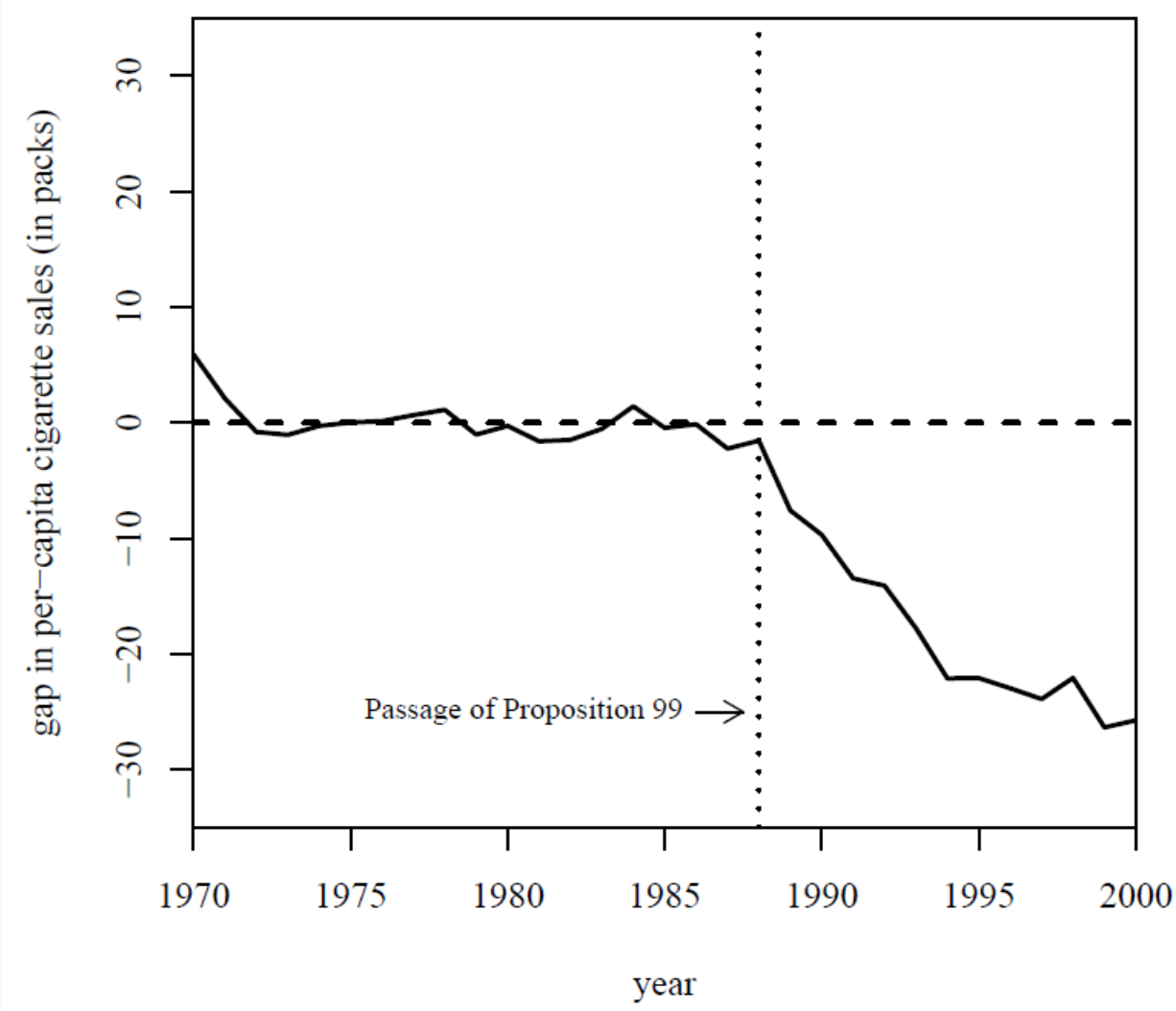
State	Weight	State	Weight
Alabama	0	Montana	0.199
Alaska	–	Nebraska	0
Arizona	–	Nevada	0.234
Arkansas	0	New Hampshire	0
Colorado	0.164	New Jersey	–
Connecticut	0.069	New Mexico	0
Delaware	0	New York	–
District of Columbia	–	North Carolina	0
Florida	–	North Dakota	0
Georgia	0	Ohio	0
Hawaii	–	Oklahoma	0
Idaho	0	Oregon	–
Illinois	0	Pennsylvania	0
Indiana	0	Rhode Island	0
Iowa	0	South Carolina	0
Kansas	0	South Dakota	0
Kentucky	0	Tennessee	0
Louisiana	0	Texas	0
Maine	0	Utah	0.334
Maryland	–	Vermont	0
Massachusetts	–	Virginia	0
Michigan	–	Washington	–
Minnesota	0	West Virginia	0
Mississippi	0	Wisconsin	0
Missouri	0	Wyoming	0

Predictor means: actual vs. synthetic California

Variables	California		Average of
	Real	Synthetic	38 control states
Ln(GDP per capita)	10.08	9.86	9.86
Percent aged 15-24	17.40	17.40	17.29
Retail price	89.42	89.41	87.27
Beer consumption per capita	24.28	24.20	23.75
Cigarette sales per capita 1988	90.10	91.62	114.20
Cigarette sales per capita 1980	120.20	120.43	136.58
Cigarette sales per capita 1975	127.10	126.99	132.81

Note: All variables except lagged cigarette sales are averaged for the 1980-1988 period (beer consumption is averaged 1984-1988).

Smoking gap between CA and synthetic CA



inference

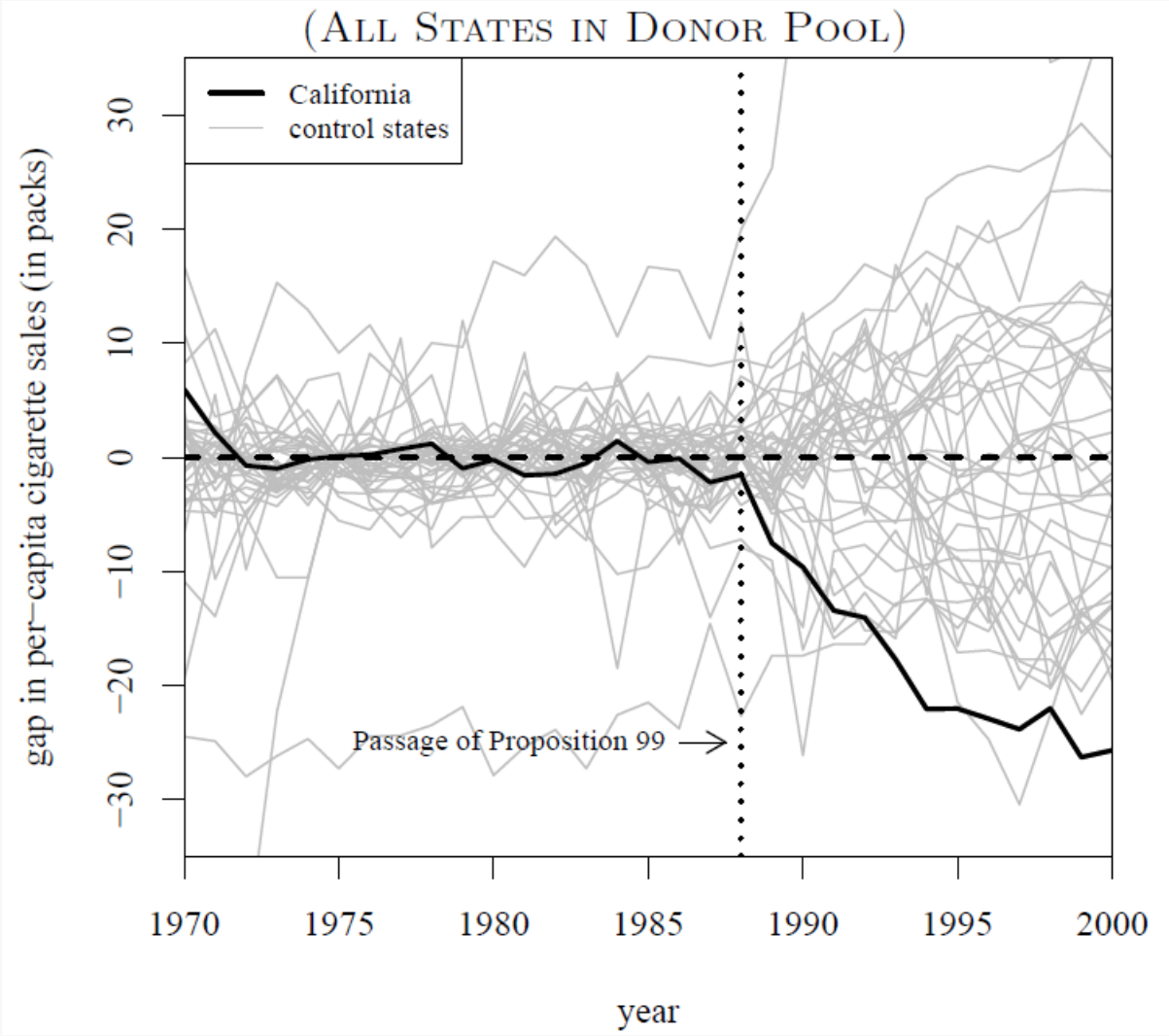
Strategy:

- Is the estimator significant?
- i.e. whether the effects estimated by the synthetic control for the treatment unit is large relative to the effect estimated for a unit chosen at random
- Valid regardless of the number of available comparison units, time periods, and whether the data are individual or aggregate

Implementation:

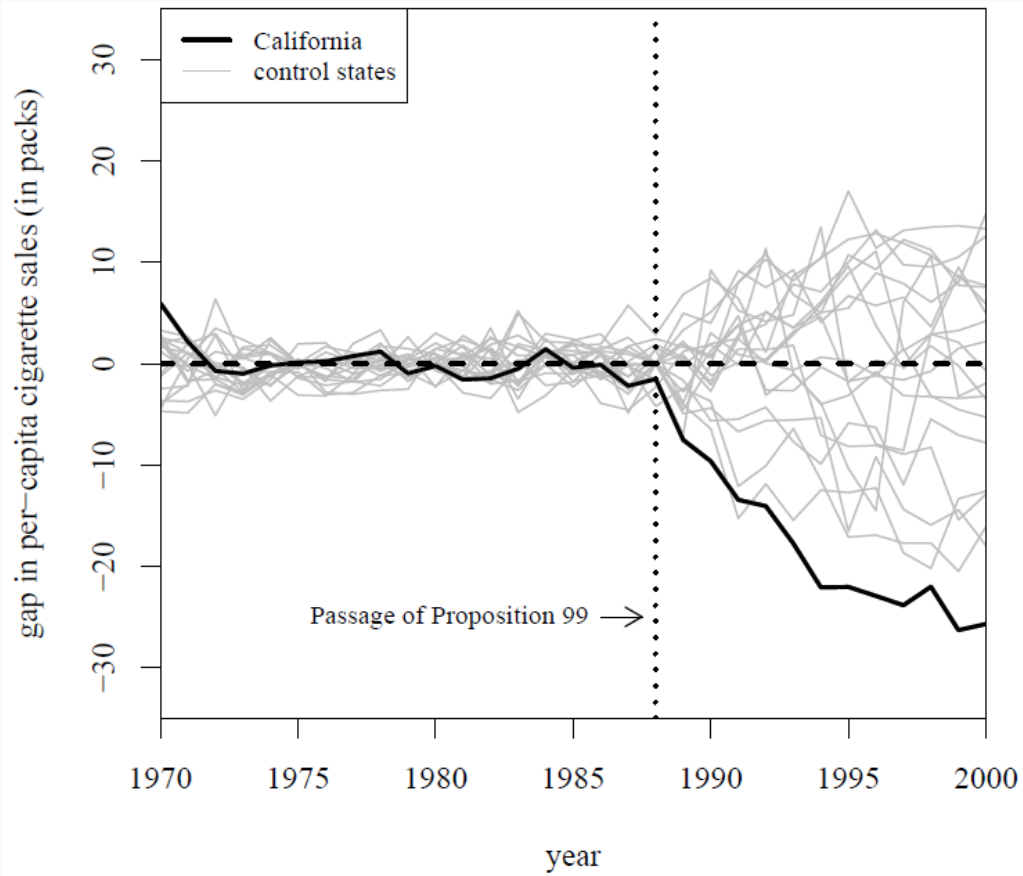
- Iteratively apply the synthetic method to each state in the donor pool and obtain a distribution of placebo effects
- Compare the gap for California to the distribution of the placebo gaps

Smoking gap for CA and 38 control states



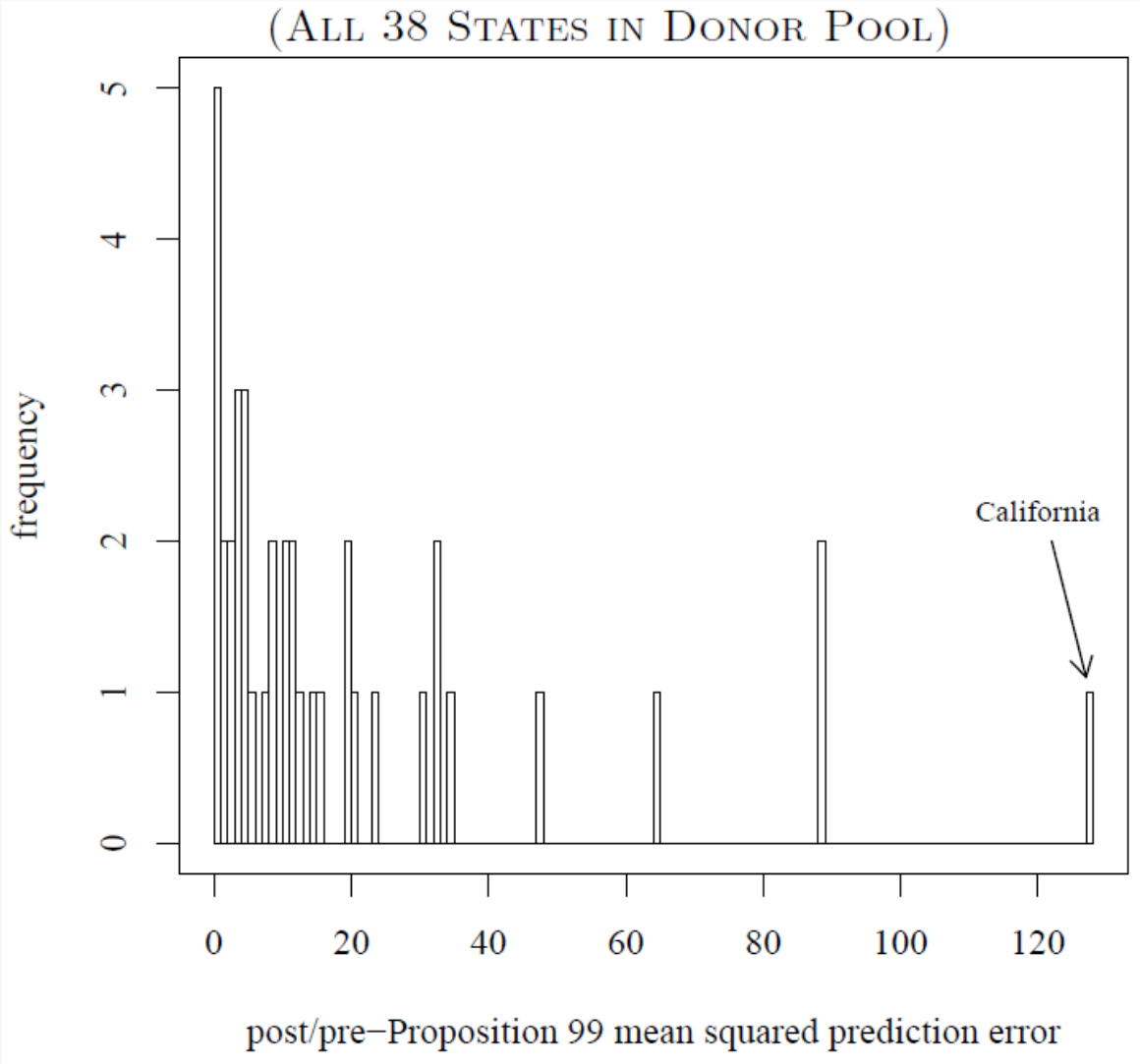
Smoking gap for CA and 19 control states

Selection: pre-policy MSPE ≤ 2 × pre-policy MSPE for CA



Estimate the significance level by comparing CA estimate with the distribution of the control group.

Ratio post-policy MSPE to pre-policy MSPE



implementation

- ssc install synth, replace

- syntax:

```
synth depvar predictorvars(x1 x2 x3) , trunit(#) trperiod(#) [ counit(numlist)
xperiod(numlist) mspeperiod() resultsperiod() nested allopt unitnames(varname)
figure keep(file) customV(numlist) optsettings ]
```

- CA Proposition 99:

```
sysuse smoking      # use the data
xtset state year    # set as panel data

synth cigsale retprice lnincome age15to24 beer cigsale(1975) cigsale(1980) ///
   cigsale(1988), trunit(3) trperiod(1989) xperiod(1980(1)1988) figure nested ///
   keep(smoking_synth)
```

dept var

indept var

Store results in 'smoking_synth'

Treated is the 3rd obs

Timing of treat

Period before treatment

Show result in figure

algorithm

Application 2: German reunification

- Synthetic control in cross-country studies
  - Cross-country regression are often criticized because they put side-by-side countries of very different characteristics.
  - The synthetic control method provides an appealing data-driven procedure to study the effects of events or interventions that take place at the country level.
  - That being said, there are also other research designs available …
    - E.g. the impact of a transaction tax policy for A-shares on stock market mechanisms.
- Application:
  - The economic impact of the 1990 German unification in West Germany.
  - Donor pool is restricted to 21 OECD countries.

Application 2: German reunification

	West Germany	Synthetic West Germany	OECD Sample excl. West Germany
GDP per-capita	8169.8	8163.1	8049.3
Trade openness	45.8	54.4	32.6
Inflation rate	3.4	4.7	7.3
Industry share	34.7	34.7	34.3
Schooling	55.5	55.6	43.8
Investment rate	27.0	27.1	25.9

Note: GDP, inflation rate, and trade openness are averaged for the 1960–1989 period. Industry share is averaged for the 1980–1989 period. Investment rate and schooling are averaged for the 1980–1985 period.

Application 2: German reunification

FIGURE 1 Trends in per Capita GDP: West Germany versus Rest of the OECD Sample

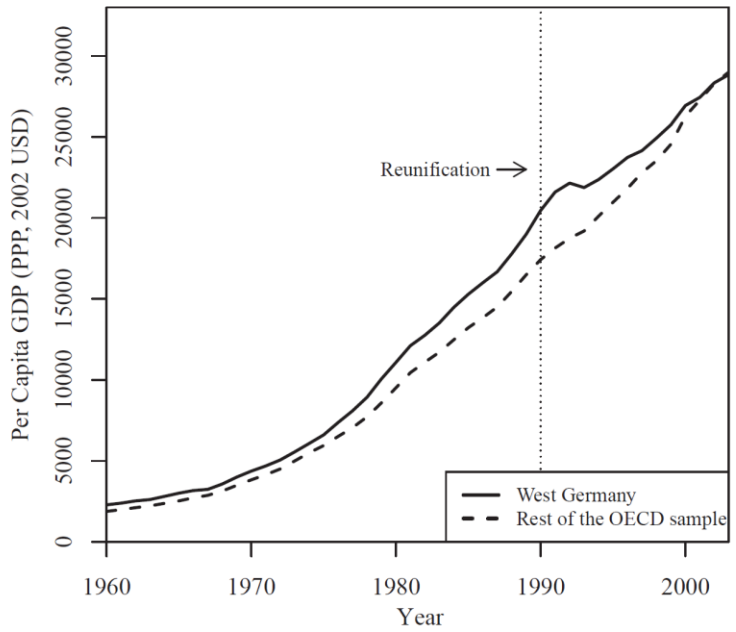
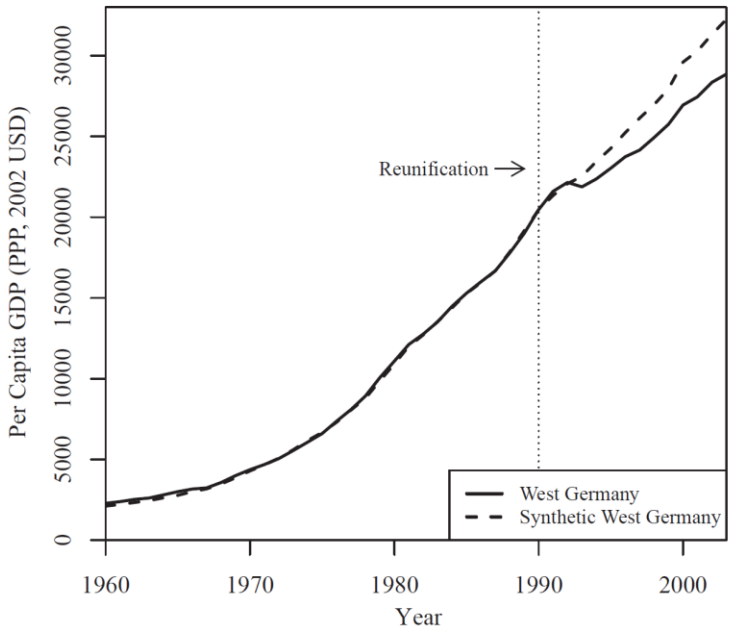
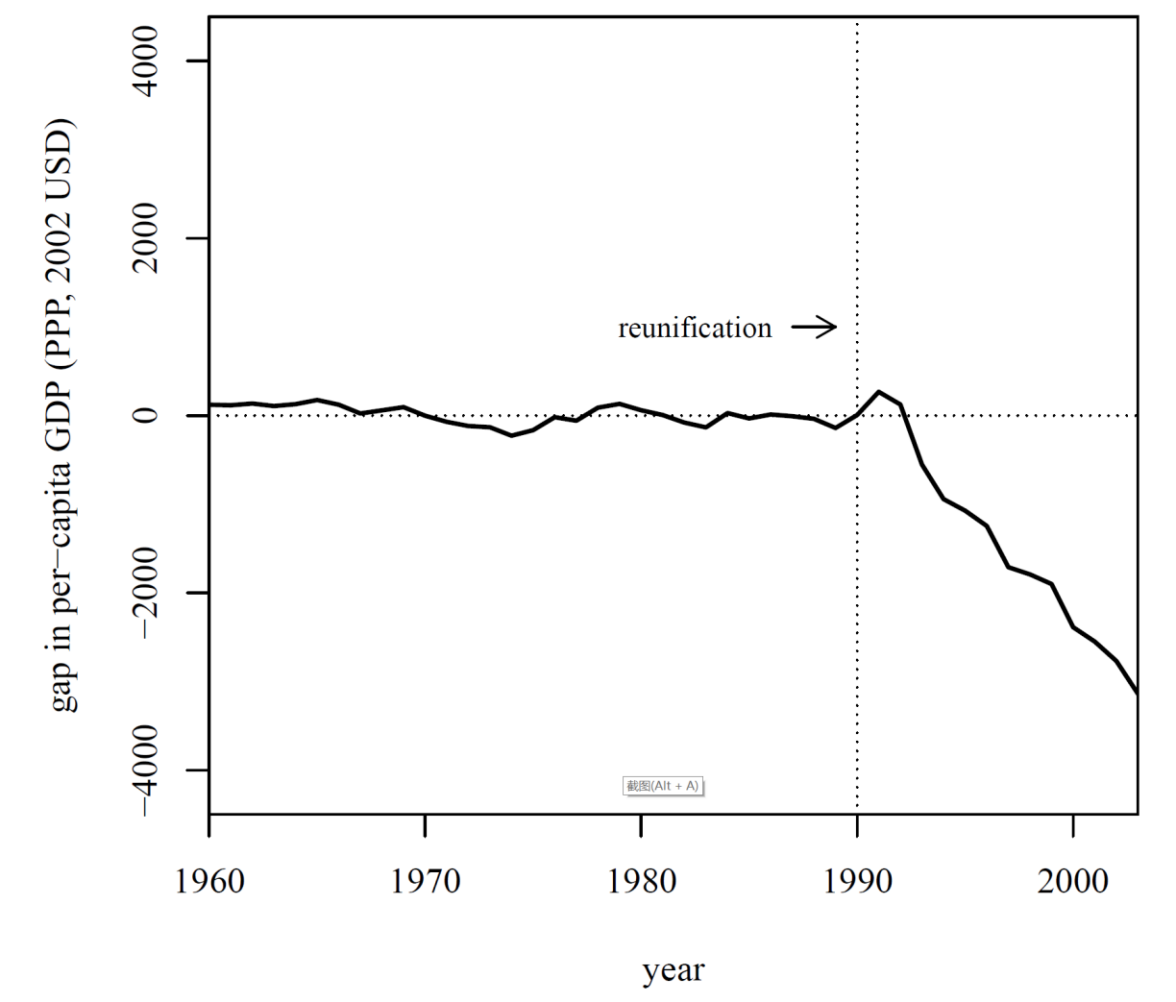


FIGURE 2 Trends in per Capita GDP: West Germany versus Synthetic West Germany



Application 2: German reunification



Application 2: German reunification

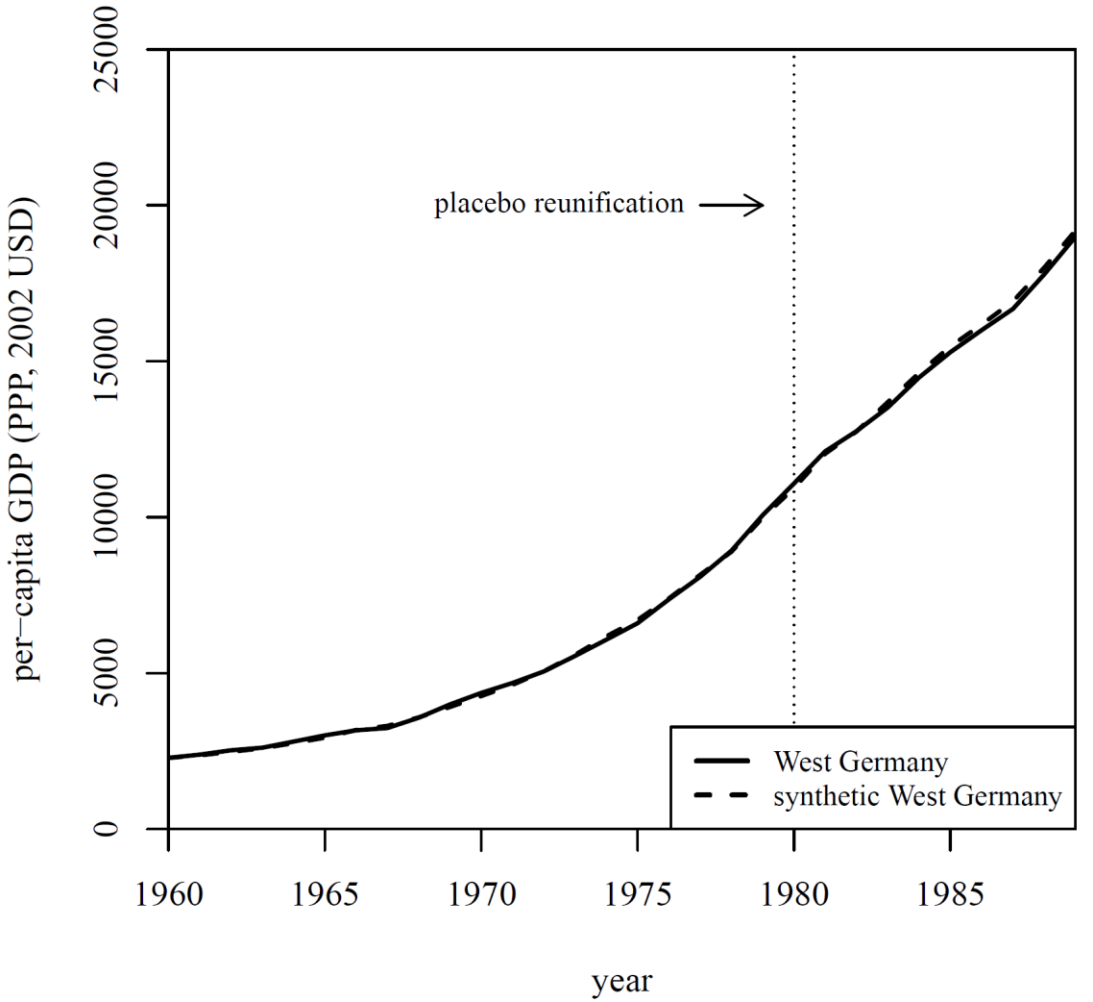
Country	Synthetic Weight	Country	Synthetic Weight
Australia	0	Netherlands	0.11
Austria	0.47	New Zealand	0
Belgium	0	Norway	0
Canada	0	Portugal	0
Denmark	0	Spain	0
France	0	Switzerland	0.17
Greece	0	United Kingdom	0
Italy	0	United States	0.14
Japan	0.11		

Application 2: German reunification

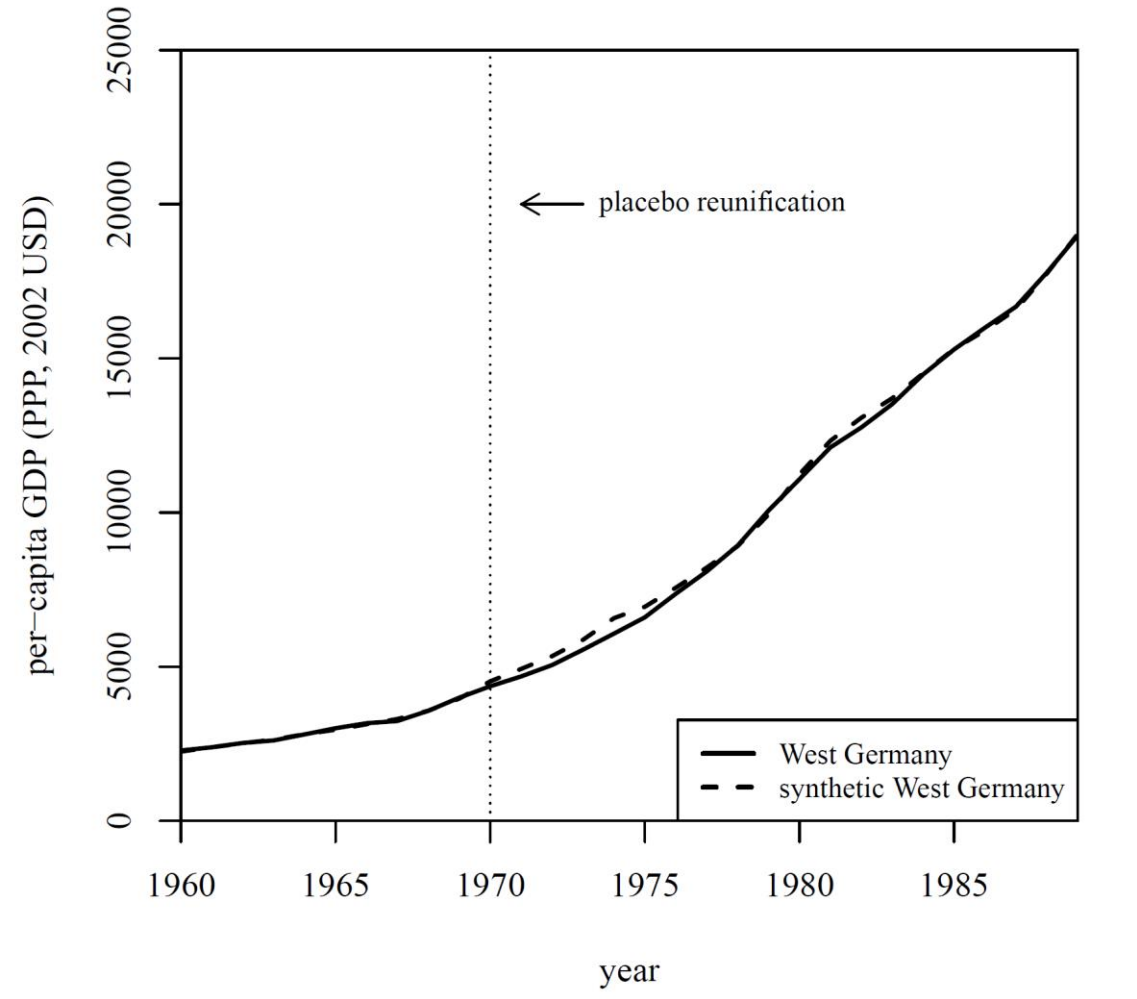
Country	Synthetic Weight	Regression Weight	Country	Synthetic Weight	Regression Weight
Australia	0	0.1	Netherlands	0.11	0.18
Austria	0.47	0.33	New Zealand	0	-0.08
Belgium	0	0.1	Norway	0	-0.07
Canada	0	0.09	Portugal	0	-0.14
Denmark	0	0.04	Spain	0	0
France	0	0.16	Switzerland	0.17	-0.06
Greece	0	0.02	United Kingdom	0	-0.04
Italy	0	-0.17	United States	0.14	0.21
Japan	0.11	0.32			

Note: Synthetic Weight: Unit weight assigned by the synthetic control method. Regression Weight: Unit weight assigned by linear regression.

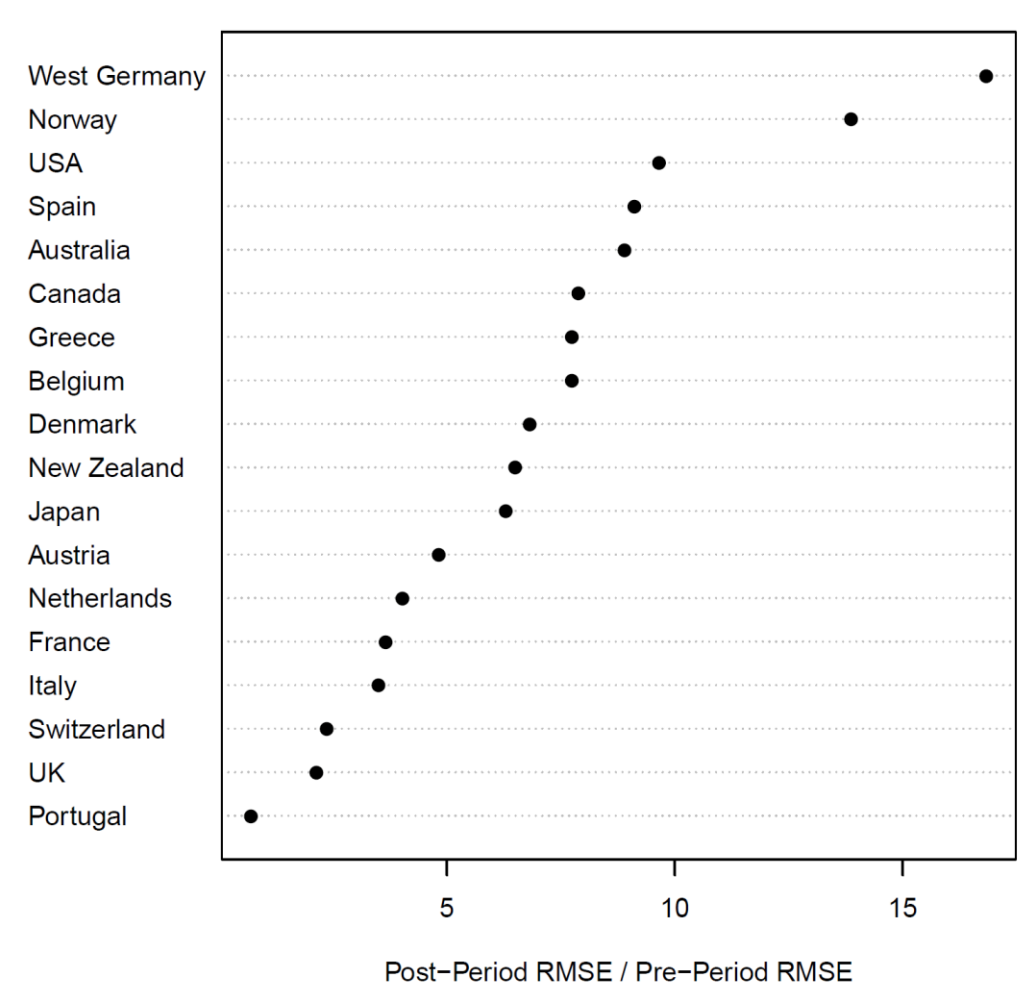
Application 2: German reunification 1980



Application 2: German reunification 1970



Application 2: German reunification



Application 3: Property tax in SH v.s. CQ

刘甲炎, 范子英. 中国房产税试点的效果评估:基于合成控制法的研究[J]. 世界经济, 2013(11):117-135.

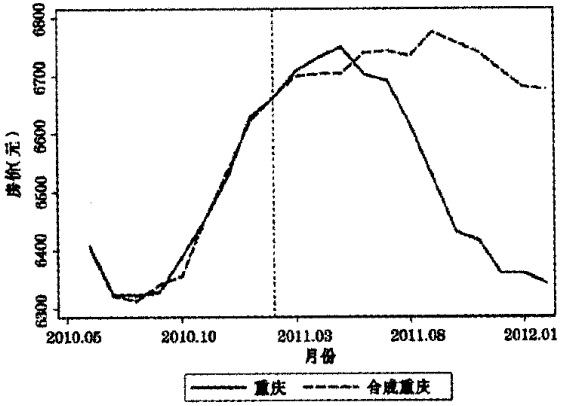


图 1　重庆实际和合成的样本房价均值

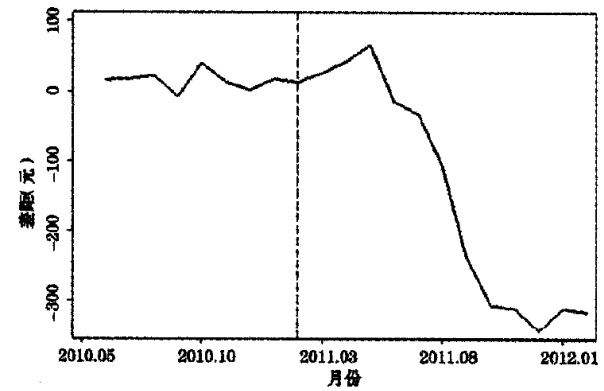


图 2　重庆实际和合成的样本房价均值差距

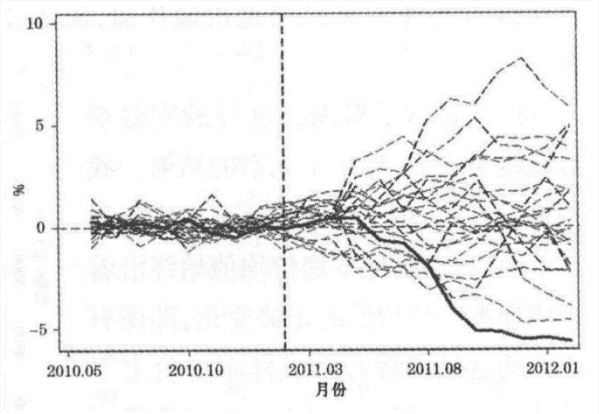


图 5　重庆和其他城市预测变动的程度分布

说明:去掉了平均标准变动程度为 0.65% 以上的城市。

Conclusion: The property tax significantly depressed housing prices in Chongqing.

Application 3: Property tax in SH v.s. CQ

CE Bai, Q Li, M Ouyang (2014) Property taxes and home prices: A tale of two cities. *Journal of Econometrics. 180*(1), 1-15

- Based on a different method
- Conclusion: housing prices increased in CQ and decreased in SH in response to the implementation of property taxes.

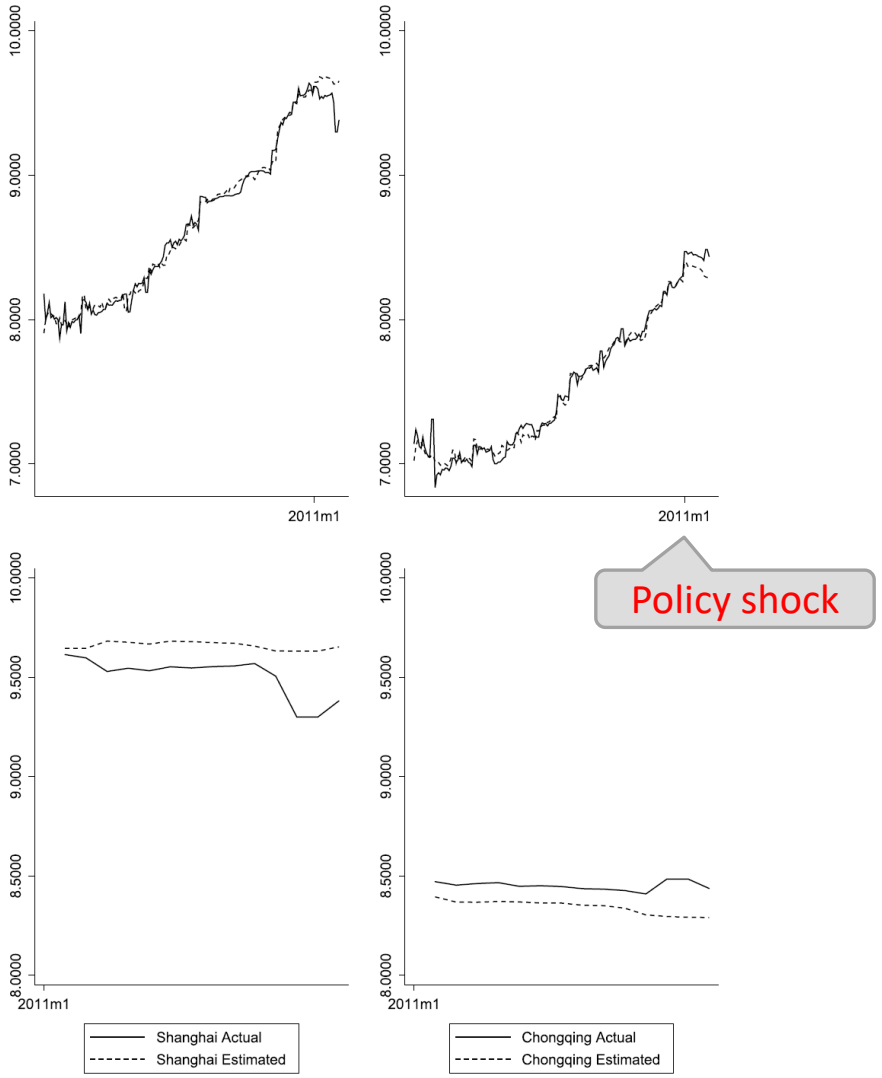


Fig. 1. Treatment effect in log levels. Note: data on actual price is from the National Development of Reform Committee (NDRC) of China; estimated price is based on data from March 1998 to January 2011. 2011m1 indicates January 2011 when the property-tax experiment is implemented. Prices are measured in log levels. The bottom panels signify the treatment effects.