Economics 513, USC, Fall 2016

Lecture 16: Synthetic Control Method

Idea

Lecture based on: 'Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program', by Abadie, Diamond and Hainmueller, JASA, 2010.

DID often applied to interventions at an aggregate level (city, state, country) using aggregate data. In the paper the intervention is the passage of Proposition 99 (increase tax on pack of cigarettes by 25c, earmarked revenues for anti-smoking campaign, spurred local ordinances that restricted smoking) and the outcome variable is the per capita number of packs per year.

Obvious comparison: California with average of other states that did not pass a similar law. See Figure 1. This obviously does not work very well.

ADH instead suggest to compare California to a weighted average of other states. If California is state 1 and we have J comparison states then estimate the effect as

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

where the weights are chosen such that the synthetic control (the weighted average 'state') resembles California in the years before the intervention. See Figure 2 and Table 2.

The estimated effect of Prop 99 is 26 packs in 2000.

Intervention effect

We want to estimate the effect of an intervention at time T_0 in state 1 with the states $2, \ldots, J+1$ used as controls. Instead of state we can use any other region or group.

Let Y_{it}^N is the outcome for state i in period t if no intervention and Y_{it}^I is the outcome for state i in period t if intervention. Note that Y_{it}^N and Y_{it}^I are defined for $t = 1, \ldots, T$ and $Y_{it}^I = Y_{it}^N$ for $t = 1, \ldots, T_0$.

Also intervention in state i has no effect on the outcome in state j. In application: passage of Prop 99 could have spread anti-smoking sentiment to control states or could have shifted ad budgets from control states to California. Downward bias in estimated effect. Cross-state purchases and smuggling could increase outcome in control states. Upward bias in estimated effect.

Intervention effect is

$$\alpha_{1t} = Y_{1t}^I - Y_{1t}^N \qquad t = T_0 + 1, \ldots, T$$

To estimate this effect we have to estimate the counterfactual outcome $Y_{1t}^N, t = T_0 + 1, \ldots, T$.

Factor model for non-treated outcomes

The DID estimator is consistent with a panel data model

$$Y_{it} = \delta_t + \beta D_{it} + \theta_t Z_i + \lambda \mu_i + \varepsilon_{it}$$

with μ_i a vector of unobserved variables. If D_{it} is correlated with μ_i we can first-difference to obtain the DID estimator (or a variant of DID).

The implied model for the non-intervention outcome is

$$Y_{it}^N = \delta_t + \theta_t Z_i + \lambda \mu_i + \varepsilon_{it}$$

The individual (here state) effect is constant over time.

ADH consider a more general model for the non-treated outcome

$$Y_{it}^N = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \varepsilon_{it}$$

This model has a time-varying individual effect, so that the trends in the outcome before the intervention do not have to be parallel.

State 1 is the treatment state and states $2, \ldots, J+1$ the control states (donor states). With nonnegative weights w_j that sum to 1, consider for $t = 1, \ldots, T_0$

$$\sum_{j=2}^{J+1} w_j Y_{jt} = \delta_t + \theta_t \sum_{j=2}^{J+1} w_j Z_j + \lambda_t \sum_{j=2}^{J+1} w_j \mu_j + \sum_{j=2}^{J+1} w_j \varepsilon_{jt}$$

so that

$$Y_{1t}^N - \sum_{j=2}^{J+1} w_j Y_{jt}^N = \theta_t (Z_1 - \sum_{j=2}^{J+1} w_j Z_j) + \lambda_t (\mu_1 - \sum_{j=2}^{J+1} w_j \mu_j) + \sum_{j=2}^{J+1} w_j (\varepsilon_{1t} - \varepsilon_{jt})$$

Organizing in vectors where in the following the subscript P indicates a T_0 vector of pre-intervention parameters or observations

$$Y_1^P - \sum_{j=2}^{J+1} w_j Y_j^P = \theta^P (Z_1 - \sum_{j=2}^{J+1} w_j Z_j) + \lambda^P (\mu_1 - \sum_{j=2}^{J+1} w_j \mu_j) + \sum_{j=2}^{J+1} w_j (\varepsilon_1^P - \varepsilon_j^P)$$

so that

$$\mu_1 - \sum_{j=2}^{J+1} w_j \mu_j = (\lambda^{P'} \lambda^P)^{-1} \lambda^{P'} \left(Y_1^P - \sum_{j=2}^{J+1} w_j Y_j^P \right) -$$

$$(\lambda^{P'} \lambda^P)^{-1} \lambda^{P'} \theta^P \left(Z_1 - \sum_{j=2}^{J+1} w_j Z_j \right) - (\lambda^{P'} \lambda^P)^{-1} \lambda^{P'} \left(\sum_{j=2}^{J+1} w_j (\varepsilon_1^P - \varepsilon_j^P) \right)$$

and upon substitution

$$Y_{1t}^N - \sum_{j=2}^{J+1} w_j Y_{jt}^N = \lambda_t (\lambda^{P'} \lambda^P)^{-1} \lambda^{P'} \left(Y_1^P - \sum_{j=2}^{J+1} w_j Y_j^P \right) +$$

$$(\theta_t - \lambda_t (\lambda^{P'} \lambda^P)^{-1} \lambda^{P'} \theta^P)(Z_1 - \sum_{j=2}^{J+1} w_j Z_j) -$$

$$\lambda_t (\lambda^{P'} \lambda^P)^{-1} \lambda^{P'} \left(\sum_{j=2}^{J+1} w_j (\varepsilon_1^P - \varepsilon_j^P) \right) + \sum_{j=2}^{J+1} w_j (\varepsilon_{1t} - \varepsilon_{jt})$$

The weighted average of the donor states is called the synthetic control. A perfect synthetic control has

$$Z_1 - \sum_{j=2}^{J+1} w_j Z_j = 0 \qquad Y_1^P - \sum_{j=2}^{J+1} w_j Y_j^P = 0$$

so that the first two terms are 0. It can be shown that the remaining terms on the right-hand side are small if T_0 is large. We can check whether the synthetic control resembles the treated state in the pre-intervention period (see figures in the paper). For a good synthetic control the expression for $Y_{1t}^N - \sum_{j=2}^{J+1} w_j Y_{jt}^N$ holds also for $t = T_0 + 1, \ldots, T$ so that we have that the counterfactual outcome for state 1 can be estimated by

$$Y_{1t}^N \approx \sum_{j=2}^{J+1} w_j Y_{jt}$$

How to choose the weights w_j? Let X_1 be the vector of Z_1 and the pre-program outcomes for 1 and X_0 the matrix where the columns have the same variables for the control states. Then choose the vector of weights W so that

$$(X_1 - X_0 W)' V (X_1 - X_0 W)$$

is minimal with V a weighting matrix.

For California see Table 1.

Inference

With aggregate data there is often no sampling uncertainty. What is the uncertainty? Remember T_0 is finite and therefore the effect of the ε_{it} is not 0. So uncertainty refers to quality of the synthetic control, i.e. how well the synthetic control approximates the pre-intervention outcomes for the intervention state.

Permutation inference: Treat each of the 38 control states as California and estimate the effect of Prop 99 (should be 0) by comparison to a synthetic control for those states. This gives us 38 estimates. See Figures 4-7.

This type of inference is also called a placebo test, because we consider the distribution of estimated effects on states that did not have the intervention

Figure 1: Trends in Per-Capita Cigarette Sales: California vs. the Rest of the United States

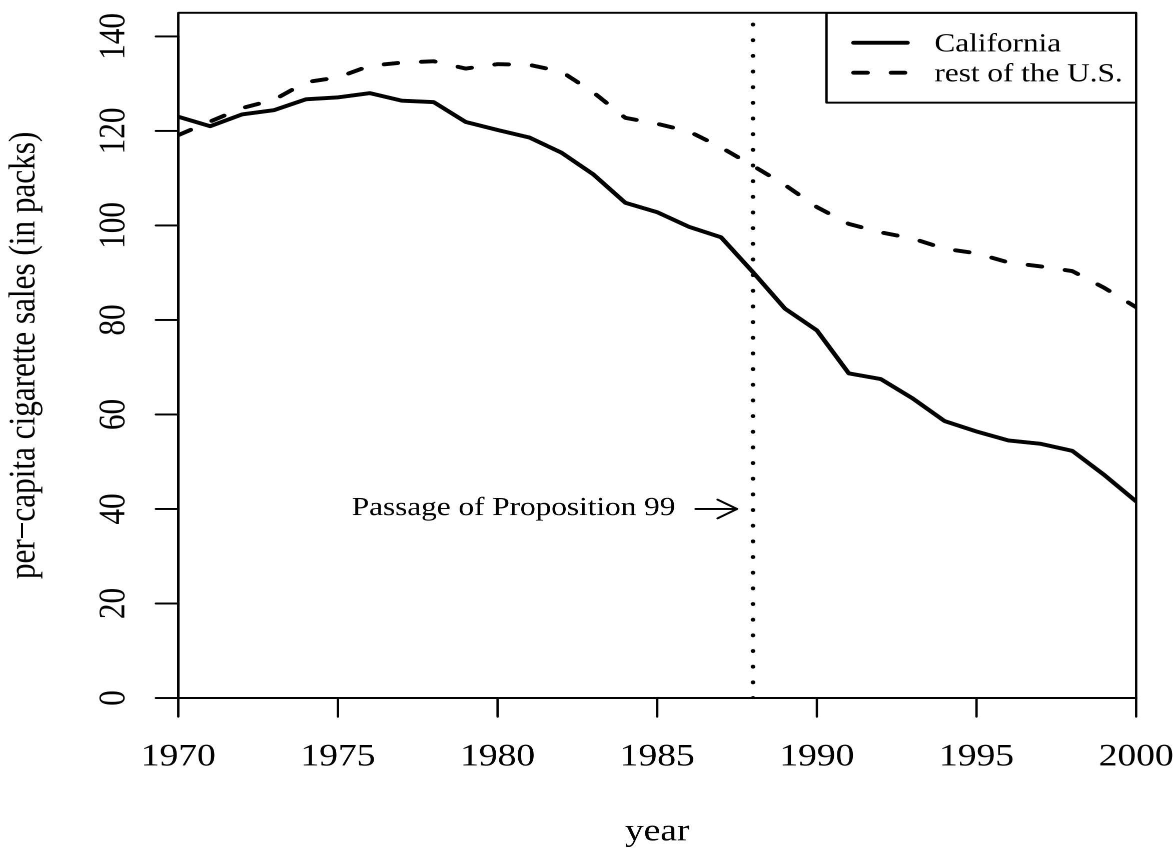

Figure 2: Trends in Per-Capita Cigarette Sales: California vs. synthetic California

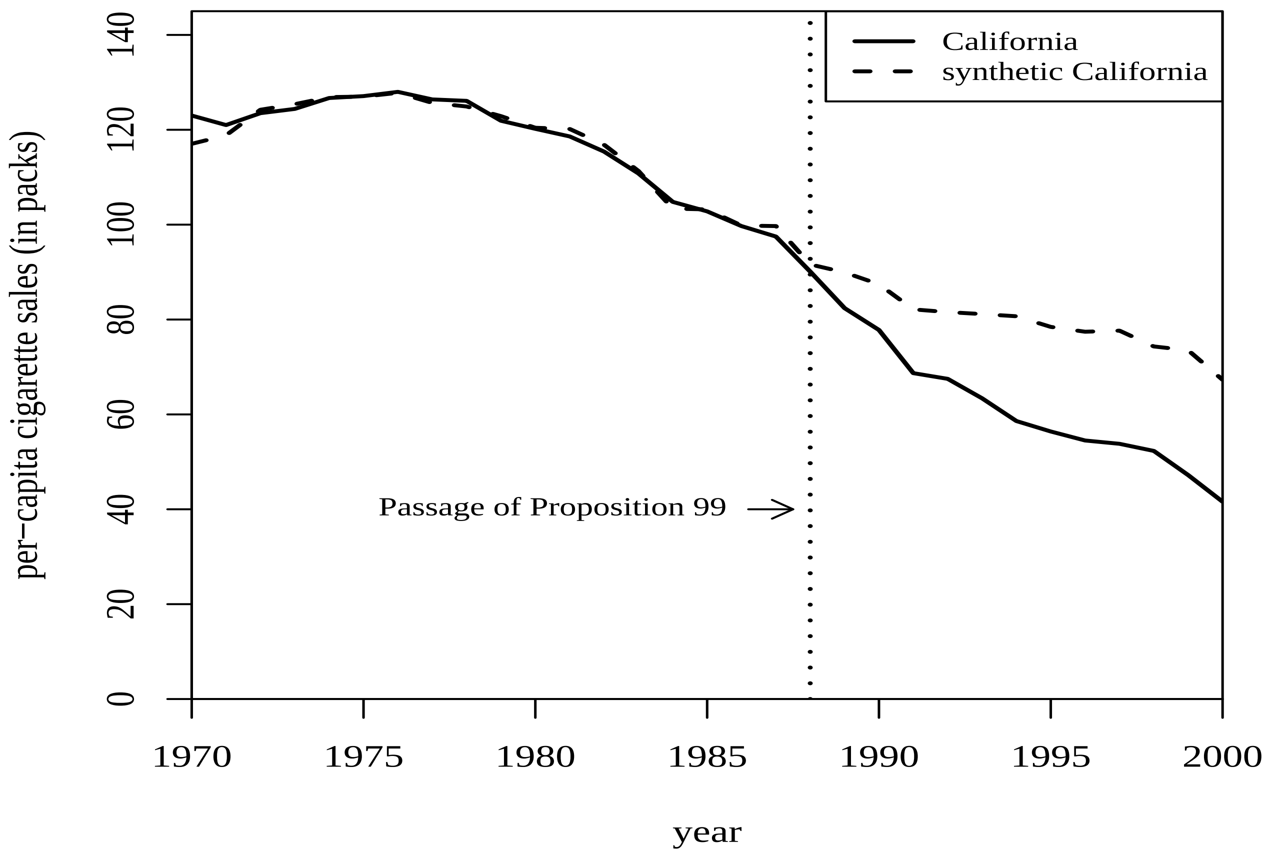

Table 2: State Weights in the Synthetic California

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

Table 1: Cigarette Sales Predictor Means

	California		Average of
Variables	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).

Figure 4: Per-Capita Cigarette Sales Gaps in California and Placebo Gaps in all 38 Control States

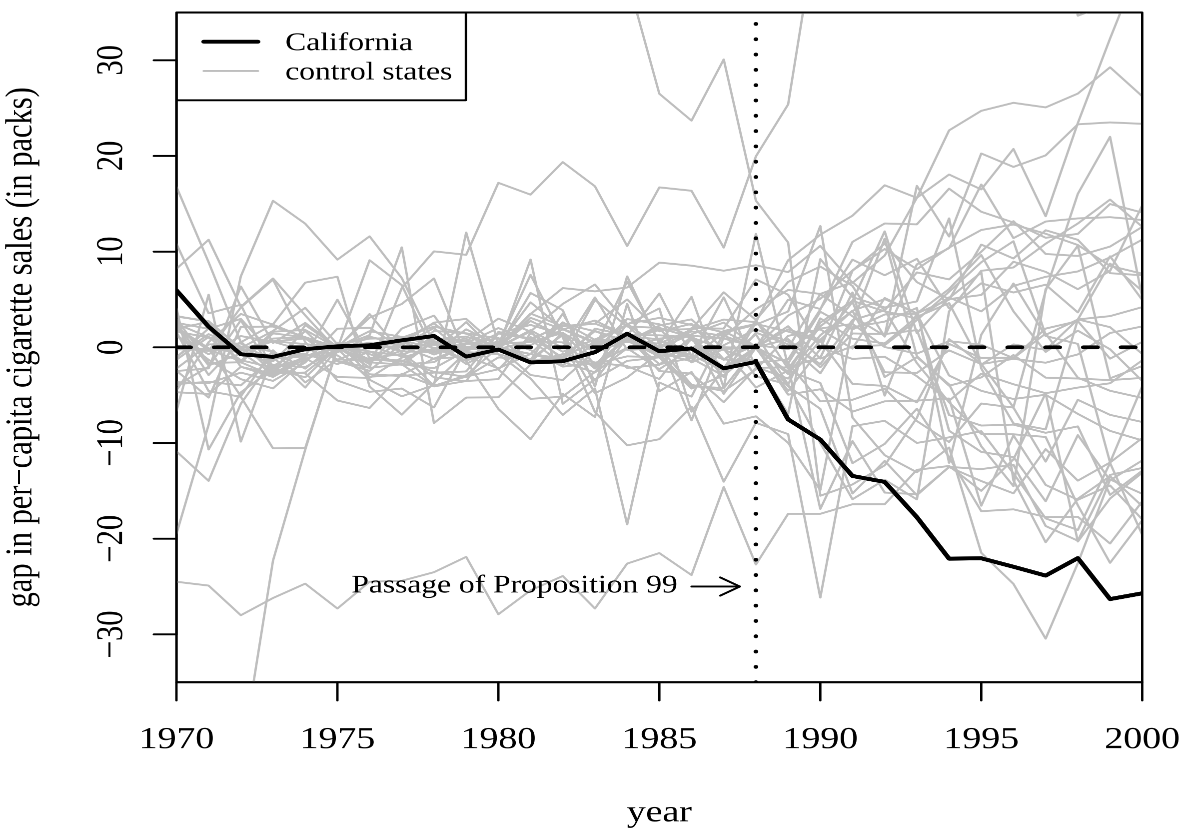

Figure 5: Per-Capita Cigarette Sales Gaps in California and Placebo Gaps in 34 Control States (Discards States with Pre-Proposition 99 MSPE Twenty Times Higher than California's)

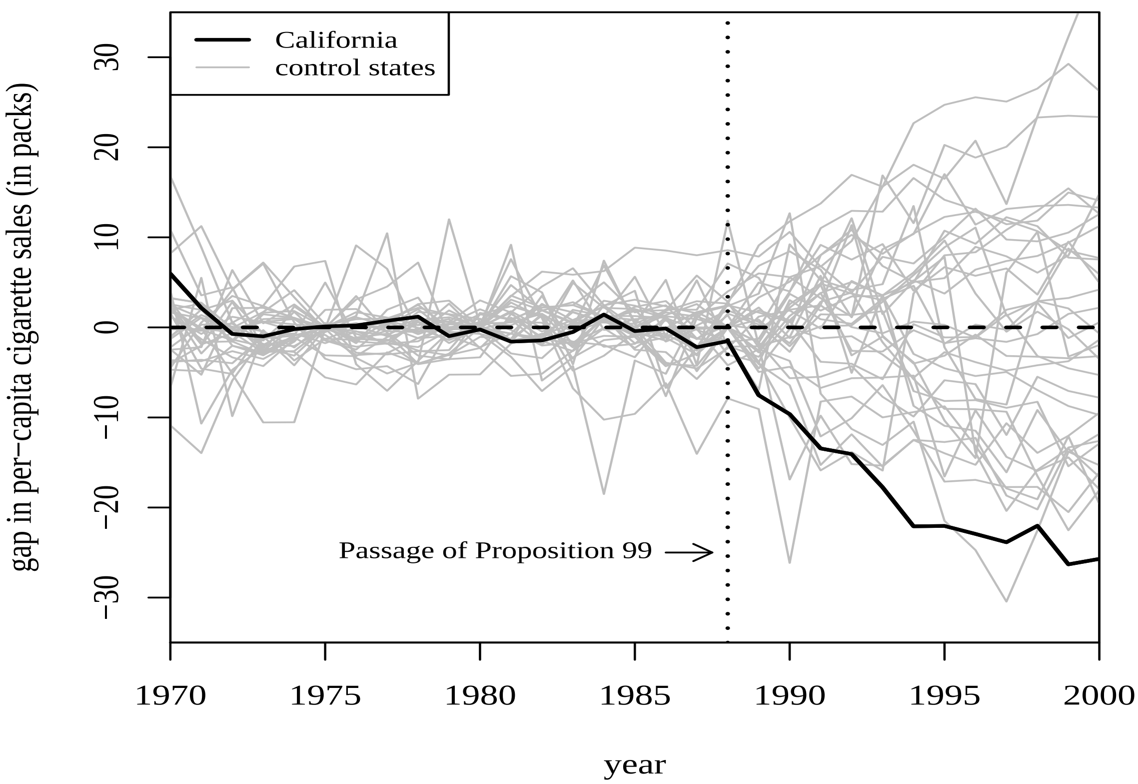

Figure 6: Per-Capita Cigarette Sales Gaps in California and Placebo Gaps in 29 Control States (Discards States with Pre-Proposition 99 MSPE Five Times Higher than California's)

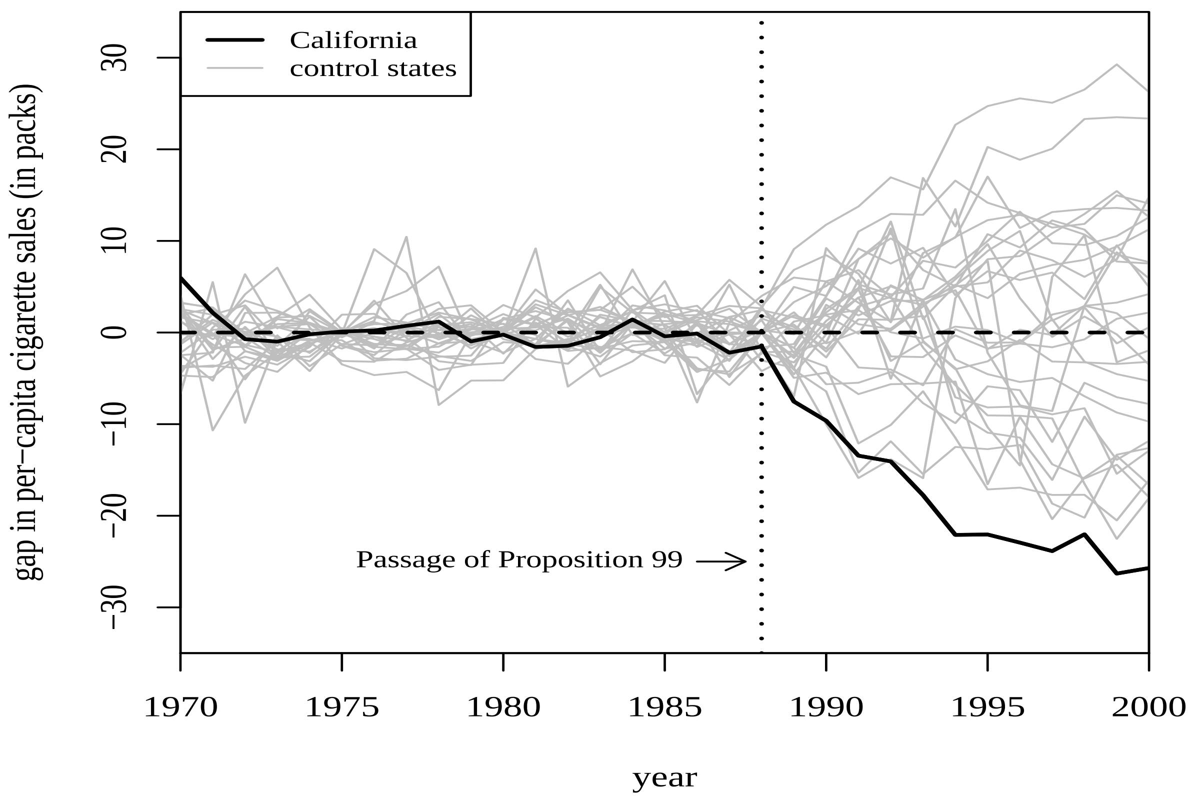

Figure 7: Per-Capita Cigarette Sales Gaps in California and Placebo Gaps in 19 Control States (Discards States with Pre-Proposition 99 MSPE Two Times Higher than California's)

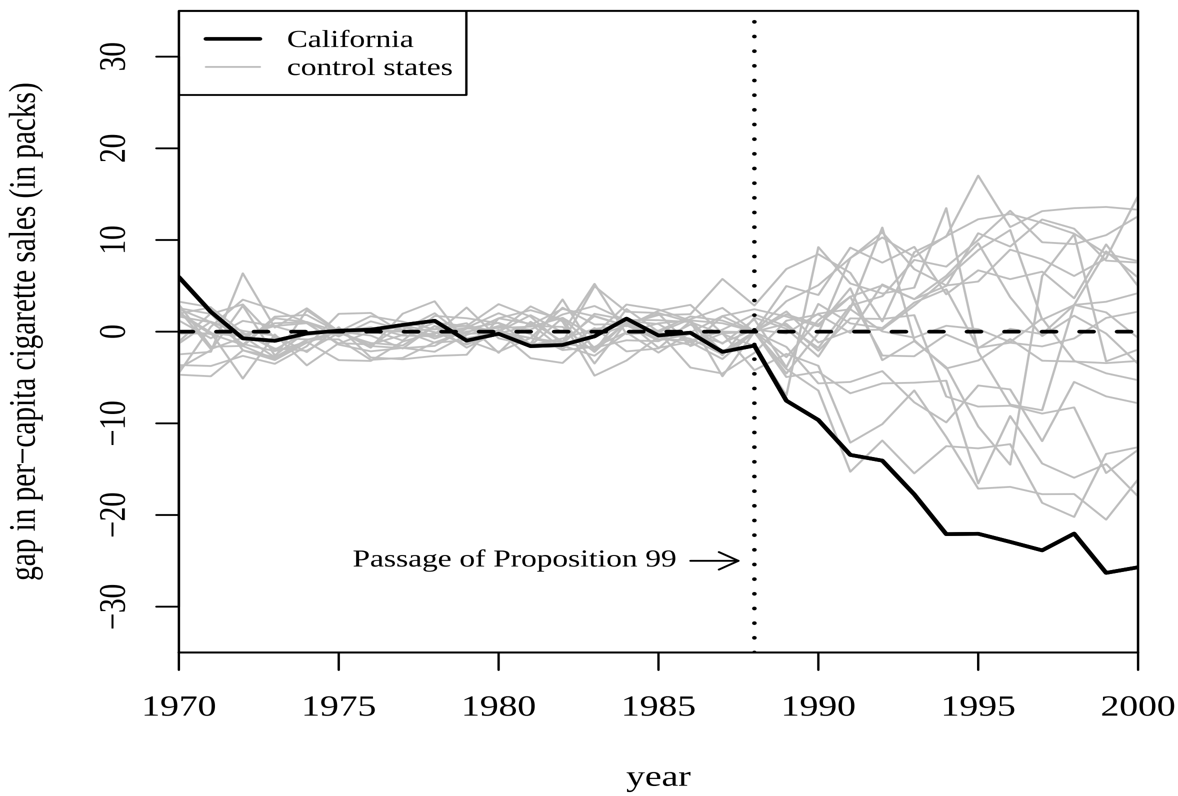

Figure A.1: Trends in Per-Capita GDP: West Germany vs. Synthetic West Germany

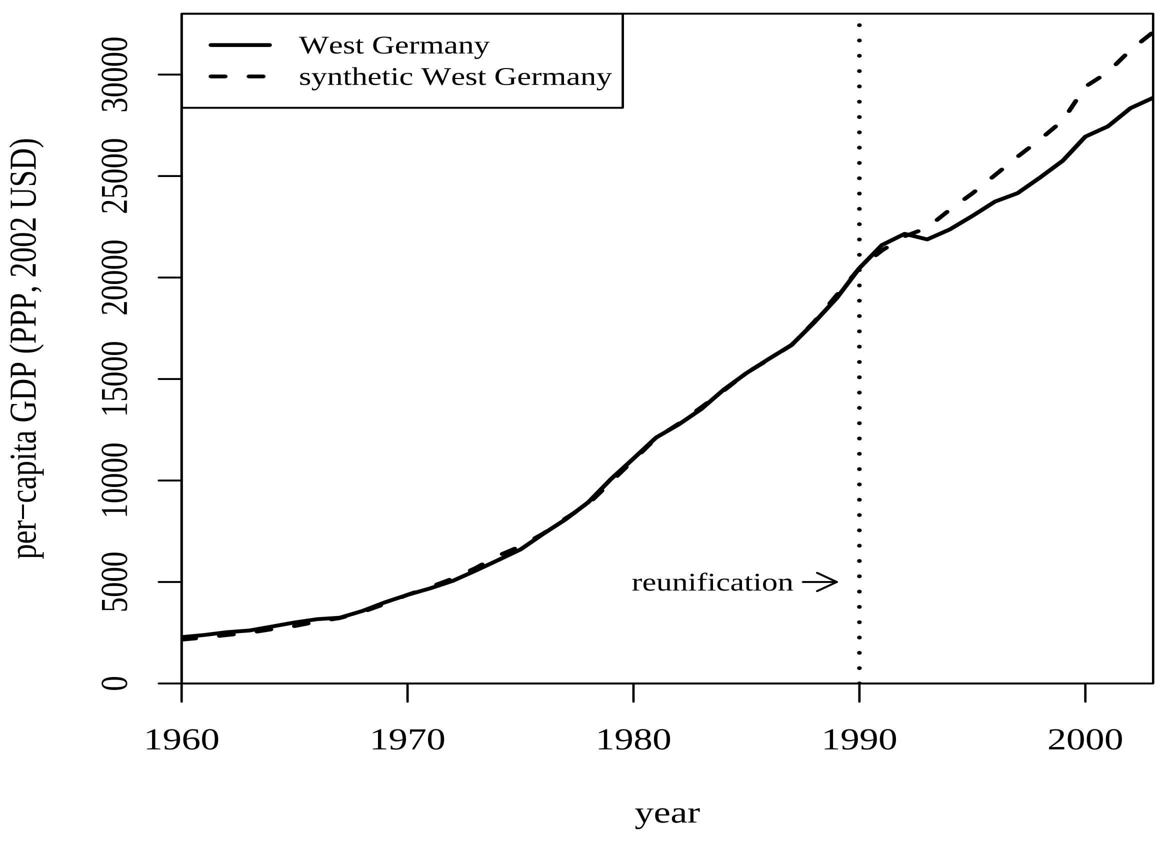

Table A.1: Economic Growth Predictor Means
before the German Reunification

	West Germany	Synthetic West Germany	OECD Sample excl. West Germany
GDP per-capita	10774.85	10808.02	10161.86
Inflation rate	3.80	5.17	9.01
Trade openness	50.59	58.41	33.47
Schooling	29.20	31.21	22.37
Investment rate	27.00	27.00	25.44
Industry share	34.69	34.64	34.28

Note: GDP, inflation rate, and trade openness are averaged for the 1970–1989 period. Industry share is averaged for the 1980–1989 period. Investment rate is averaged for the 1980–1985 period. Schooling is from year 1985.

Table A.2: Country Weights in the Synthetic West Germany

Country	Weight	Country	Weight
Australia	0	Japan	0.127
Austria	0.421	Netherlands	0.137
Belgium	0	New Zealand	0
Canada	0	Norway	0
Denmark	0	Portugal	0
Finland	0	Spain	0
France	0	Sweden	0
Greece	0	Switzerland	0.153
Ireland	0	United Kingdom	0
Italy	0	United States	0.161