

From the research bench

What is the synthetic control method?

A synthetic control builds the missing comparison for one treated unit from a weighted mix of untreated ones. It only works if the mix fits the past.

Author	Published	Reading time	Words
Sinan Isoglu, MBA (Quantic)	18 October 2026	10 min	2,238
Commercial growth leader, lecturer and doctoral researcher			

Management summary The synthetic control method estimates the effect of an event or intervention on one aggregate unit, such as a state, a region or a market, by comparing it with a weighted average of untreated units chosen to match it before the intervention. This concept guide sets out how the weights are chosen, using the authors' own example of California's 1988 tobacco program (Abadie et al.), what the method assumes compared with difference-in-differences (Abadie and Cattaneo), how placebo runs test whether an effect stands out, when the method should not be used, and gives a one-page record for a synthetic control study. The sources are a methods paper and a review.

Contents	§ 1	What is the synthetic control method?	2
	§ 2	How is a synthetic control built?	3
	Figure 1	The five states that make up synthetic California	4
	§ 3	Why not use one comparison unit, or the average of all of them?	4
	Figure 2	California before Proposition 99: the real state, its synthetic version and the donor average	5
	§ 4	What does the synthetic control method assume?	6
	§ 5	How do you test whether the gap could be chance?	7
	§ 6	How do you write a synthetic control study down?	8
	Figure 3	A synthetic control study record	8
	§ 7	Where does the evidence on synthetic control stop?	9
		References	10

What is the synthetic control method?

§ 1 of 7

The synthetic control method estimates what an intervention did to one unit by building that unit's missing comparison from other units. Abadie and Cattaneo (2018)¹ define the comparison: "A synthetic control is a weighted average of untreated units chosen to reproduce characteristics of the treated unit before the intervention." The estimated effect in each period after the intervention is the gap between the treated unit's outcome and the synthetic control's.

¹ Abadie, A., & Cattaneo, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics*, 10, 465-503.

The method was built for a common situation. Abadie et al. (2010)² note that "many policy interventions and events of interest in social sciences take place at an aggregate level (countries, regions, cities, etc.) and affect a small number of aggregate units". Comparing such a unit with others is a comparative case study, and the authors name two weak points: how the comparison is chosen and how uncertainty is measured. On the first: "Researchers often select comparison groups on the basis of subjective measures of affinity between affected and unaffected units." The synthetic control narrows that choice rather than removing it: data-driven procedures "reduce discretion in the choice of the comparison control units", while the donor pool and the predictors remain the researcher's to set. "The idea behind the synthetic control approach is that a combination of units often provides a better comparison for the unit exposed to the intervention than any single unit alone."

² Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490), 493-505.

In this guide, a synthetic control study is an estimate of one unit's counterfactual path from a stated pool of untreated units, with weights that are visible, a pre-intervention fit that is reported, and a placebo test that says whether the post-intervention gap stands out. Commercially, the same shape appears when one country gets a new price, one region a campaign or one territory a new sales model, and no randomized comparison exists; that application is the author's, not the sources'.

The [incrementality guide](#) lists synthetic control regions among market-level tests of what marketing spend adds. This page sets out

the conditions such a test has to meet before its result can carry that weight.

How is a synthetic control built?

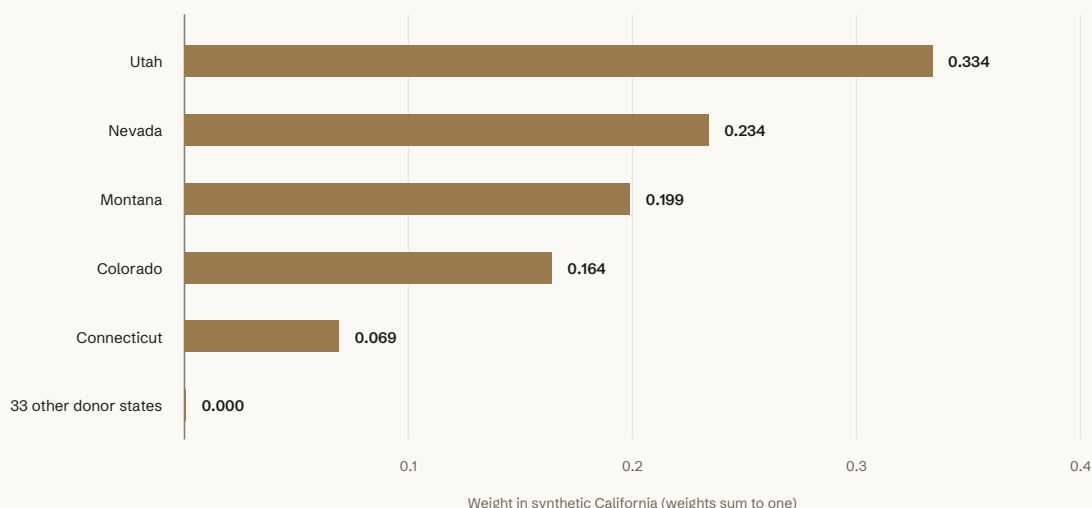
§ 2 of 7

From a donor pool, with weights that are nonnegative and sum to one. Abadie and Cattaneo (2018)¹ describe the construction: the untreated units form “an untreated reservoir of potential controls (a donor pool)”, and the weights are chosen to bring the synthetic control as close as possible to the treated unit on characteristics measured before the intervention. In Abadie et al. (2010)² those characteristics include earlier values of the outcome, and in their application the authors chose the relative importance of the characteristics “such that the mean squared prediction error of the outcome variable is minimized for the preintervention periods”.

Their example is California’s Proposition 99, “a large-scale tobacco control program that California implemented in 1988”. The donor pool held 38 states after the authors removed states that ran their own large tobacco programs and states that raised cigarette taxes by 50 cents or more between 1989 and 2000. The result was a short list: “smoking trends in California prior to the passage of Proposition 99 is best reproduced by a combination of Colorado, Connecticut, Montana, Nevada, and Utah. All other states in the donor pool are assigned zero W-weights.”

Figure 1 The five states that make up synthetic California

Five of 38 donor states carry all of the weight. The method states the combination openly, so a reader can judge whether it is a plausible stand-in.



The weights come from matching California up to 1988, before Proposition 99 took effect, on cigarette sales and their predictors, not from any outcome after the program began.

Source Abadie et al. (2010), Table 2, p. 500. Weights of each state in the donor pool of 38 states; the other 33 states carry a weight of zero. The authors' estimate, not a re-estimation.

Few nonzero weights are what the method generally returns when the treated unit cannot be reproduced exactly. Abadie and Cattaneo (2018)¹ note that, ruling out degenerate cases, if the treated unit's pre-intervention characteristics lie outside the convex hull of the donor units' characteristics, the weights are "unique and sparse", meaning the weight vector "has only a few nonzero elements". Synthetic California is consistent with that: its fit before the program is close but not exact (Figure 2).

Why not use one comparison unit, or the average of all of them?

§ 3 of 7

Because neither may resemble the treated unit before the intervention. In the tobacco example, the population-weighted average of the 38 donor states did not. In the authors' words, it "does not seem to provide a suitable control group for California". The synthetic version did.

Figure 2 California before Proposition 99: the real state, its synthetic version and the donor average

The synthetic California tracks the real one on past cigarette sales. The population-weighted average of the 38 donor states does not.

Predictor	California	Synthetic California	Population-weighted average of 38 donor states
Cigarette sales per person, 1988 (packs)	90.10	91.62	114.20
Cigarette sales per person, 1980 (packs)	120.20	120.43	136.58
Cigarette sales per person, 1975 (packs)	127.10	126.99	132.81
Retail price (cents)	89.42	89.41	87.27
Beer consumption per person (gallons)	24.28	24.20	23.75
Percent aged 15–24	17.40	17.40	17.29
Ln(GDP per capita)	10.08	9.86	9.86

The synthetic version misses on income, which the authors found added little to predicting cigarette sales once the other variables were in.

Source Abadie et al. (2010), Table 1, p. 499. Predictor means; all variables except lagged cigarette sales are averaged for 1980–1988 (beer consumption for 1984–1988). Cigarette sales in packs per person, retail price in cents, beer consumption in gallons per person. GDP per capita is measured in 1997 dollars; the authors' appendix describes the variable as logged per-capita state personal income. The third column is the population-weighted average of the 38 donor states (p. 499).

Abadie et al. (2010)² name two features that set the method apart from regression. “Relative to traditional regression methods, transparency and safeguard against extrapolation are two attractive features of the synthetic control method.” The weights show “the relative contribution of each control unit to the counterfactual of interest”, and because they can be restricted to be nonnegative and to sum to one, the method “provides a safeguard against extrapolation”. A further point, which the authors tie to research honesty in observational studies: “because the choice of a synthetic control does not require access to

postintervention outcomes, the synthetic control method allows researchers to decide on study design without knowing how those decisions will affect the conclusions of their studies”.

What does the synthetic control method assume?

§ 4 of 7

Less than difference-in-differences assumes about how unobserved factors act over time, but it asks for two things difference-in-differences does not: donors whose combination can closely reproduce the treated unit before the intervention, and, for the authors’ argument under their factor model that the estimate is close to unbiased, a long pre-intervention record. Difference-in-differences assumes common trends; in Abadie and Cattaneo’s (2018)¹ words, “in the absence of the treatment, the average outcome for the treated and the average outcome for the nontreated would have experienced the same variation over time”, which the [parallel trends guide](#) treats in full. The model behind the synthetic control, Abadie et al. (2010)² write, “extends the traditional linear panel data (difference-in-differences) framework, allowing that the effects of unobserved variables on the outcome vary with time”.

Spelled out, the method comes with three conditions. The first is fit. Where no combination of donors resembles the treated unit, “the fit may be poor and then we would not recommend using a synthetic control”, and even with a good fit “interpolation biases may be large” when the outcome is not linear in the units’ characteristics and the units differ widely, which the authors suggest limiting by restricting “the donor pool to regions with similar characteristics”. Under their factor model, the authors’ argument that the estimate is close to unbiased needs both a close fit and a long pre-intervention record: given weights that reproduce the treated unit before the intervention, the bias is close to zero “if the number of preintervention periods is large relative to the scale of the transitory shocks”. Under an autoregressive model they show that a single pre-intervention period can be enough, so how long the record must be depends on how the outcome is generated.

The second is no interference. The authors’ notation carries “the usual assumption of no interference between units”: the untreated units’ outcomes do not react to the treated unit’s intervention. In the tobacco case, the authors named as “Perhaps the most important concern” that anti-tobacco sentiment from California “could have spread to other states, contaminating the donor pool”. They judged such violations not severe, and noted that this one would, if anything, have made their

estimate smaller. Commercially, a campaign in one region that reaches buyers in another is the same problem, which the interference guide treats in general.

The third is that the start date is right. “In practice, interventions may have an impact prior to their implementation (e.g., via anticipation effects)”, and then the start can be moved to “the first period in which the outcome may possibly react to the intervention”.

How do you test whether the gap could be chance?

§ 5 of 7

By running the same method on units that were not treated and seeing whether the treated unit’s gap stands out. In the tobacco study the gap was large: the authors estimate “that by the year 2000 annual per-capita cigarette sales in California were about 26 packs lower than what they would have been in the absence of Proposition 99”. The question is whether a gap like California’s could arise by chance. Abadie et al. (2010)² ask: “How often would we obtain results of this magnitude if we had chosen a state at random for the study instead of California?” They reassign the intervention to each of the 38 donor states in turn, estimate a placebo gap for each, and compare. One summary is the ratio of the mean squared prediction error after the intervention to the same error before it, which “obviates choosing a cutoff for the exclusion of ill-fitting placebo runs”. California’s ratio was the largest of the 39. If the intervention were assigned at random, the probability “of obtaining a post/pre-Proposition 99 MSPE ratio as large as California’s is $1/39 = 0.026$ ”. That probability is a rank, not a margin of error around the 26 packs. With 38 donor states, $1/39$ is the smallest value this permutation inference can return; when the authors instead compare the gaps of only the 19 donor states they can fit almost as well as California, the same reasoning gives 5 percent, again the smallest value 20 units allow.

The test answers a narrower question than a randomized experiment would. The authors describe it as inference that “examines whether or not the estimated effect of the actual intervention is large relative to the distribution of the effects estimated for the regions not exposed to the intervention”. This is “informative inference if under the hypothesis of no intervention effect the estimated effect of the intervention is not expected to be abnormal relative to the distribution of the placebo effects”. It becomes classical randomization inference only “for the case where the intervention is indeed randomized across regions, a rather restrictive condition”. Abadie and Cattaneo (2018)¹ call it “an

inferential method for synthetic controls based on Fisher's randomization inference".

How do you write a synthetic control study down?

§ 6 of 7

As one record that fixes the treated unit, the start date, the donor pool and its exclusions, the predictors and the pre-intervention fit before any outcome after the start is looked at, and then adds the placebo results and the statement they support. The order follows the authors' point that the design can be settled without the post-intervention outcomes.

Figure 3 A synthetic control study record

Fix the unit, the start, the donor pool, the predictors and the fit before reading the outcome. Then record the placebo runs and the statement they support.

Unit and start	Donor pool	Predictors and fit	Interference	Placebo runs	Supported statement
Which unit was treated, and when could it first react?	Which units, and which were excluded for similar changes?	Which pre-period values, and how close is the fit?	Could the change reach any donor unit?	Where does the post/pre error ratio rank among placebo runs?	Effect, no clear effect, or no usable fit?

If the pre-period fit is poor, the record ends there. A gap after a poor fit is not an estimate.

One row per study. Fill the first three fields before any outcome after the start date is opened.

Source Author's decision worksheet. The fields follow Abadie et al. (2010), pp. 494–503, on the donor pool, predictors, fit, interference, anticipation and placebo inference, and Abadie and Cattaneo (2018), pp. 483 and 485. The fields are reader inputs; no company data is supplied.

An illustration, with no company data: a firm raises list prices in one country and wants the effect on unit sales. The record names the country and the announcement date, since buyers may react to the announcement before the new prices apply; it lists the other countries as the donor pool and removes any that changed prices or ran a major

campaign in the same months; it fixes the predictors, such as several years of monthly sales; and it checks whether buyers in the treated country could buy through a neighboring one. If the fit before the change is poor, the record stops there.

Where does the evidence on synthetic control stop? § 7 of 7

At method and one application. Abadie et al. (2010)² is a methods paper whose application is a state-level public-health program, with an outcome measured yearly over three decades. Abadie and Cattaneo (2018) is a review that lists variants for cases the original does not handle well: with many treated units “there may be synthetic controls that are not unique or sparse”, and other authors “propose variants of the synthetic control framework that extrapolate outside of the support” of the donor units. Neither source gives a minimum number of donor units or pre-intervention periods, and neither tests the method on commercial data. Commercial settings often have short histories, few comparable markets and units that influence each other, which makes the fit, the long pre-intervention record and the no-interference condition harder to meet; that reading, and every commercial example in this guide, is the author’s.

 End of essay

References

- 1 Abadie, A., & Cattaneo, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics*, 10, 465-503.
<https://doi.org/10.1146/annurev-economics-080217-053402>
- 2 Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490), 493-505.
<https://doi.org/10.1198/jasa.2009.ap08746>

How to cite this essay

Isoglu, S. (2026, October 18). What is the synthetic control method? Sinan Isoglu.
<https://isoglu.com/insights/journal/what-is-the-synthetic-control-method/>
Online version:
<https://isoglu.com/insights/journal/what-is-the-synthetic-control-method/>

Sinan Isoglu, MBA (Quantic)

Commercial growth leader, lecturer and doctoral researcher
© 2026 Sinan Isoglu · isoglu.com · All rights reserved