The Stata Journal (2023)
23, Number 3, pp. 597–624 DOI: 10.1177/1536867X231195278

synth2: Synthetic control method with placebo tests, robustness test, and visualization

Guanpeng Yan
Shandong University
Jinan, China
guanpengyan@yeah.net

Qiang Chen
Shandong University
Jinan, China
qiang2chen2@126.com

Abstract. The synthetic control method (Abadie and Gardeazabal, 2003, *American Economic Review* 93: 113–132, Abadie, Diamond, and Hainmueller, 2010, *Journal of the American Statistical Association* 105: 493–505) is a popular method for causal inference in panel data with one treated unit that often uses placebo tests for statistical inference. While the synthetic control method can be implemented by the excellent command `synth` (Abadie, Diamond, and Hainmueller, 2011, Statistical Software Components S457334, Department of Economics, Boston College), it is still inconvenient for users to conduct placebo tests. As a wrapper program for `synth`, our proposed `synth2` command provides convenient utilities to automate both in-space and in-time placebo tests, as well as the leave-one-out robustness test. Moreover, `synth2` produces a complete set of graphs to visualize covariate or unit weights, covariate balance, actual or predicted outcomes, treatment effects, placebo tests, ratio of posttreatment mean squared prediction error to pretreatment mean squared prediction error, pointwise p-values (two-sided, right-sided, and left-sided), and the leave-one-out robustness test. We illustrate the use of the `synth2` command by revisiting the classic example of California's tobacco control program (Abadie, Diamond, and Hainmueller 2010).

Keywords: st0722, synth2, synth, synthetic control method, placebo test, robustness test, causal inference

1 Introduction

The synthetic control method (SCM) (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010) is a widely used approach for causal inference in panel data with one treated unit. Hailed as "arguably the most important innovation in the policy evaluation literature in the last 15 years" (Athey and Imbens 2017), the SCM has spawned a large literature; see Abadie (2021) for an excellent review. Basically, for a treated unit, the SCM constructs its counterfactual outcomes via a linear combination of untreated units with optimal weights constrained to be nonnegative and summed to one. For each posttreatment period, the treatment effect is then estimated as the difference between the observed and counterfactual outcomes for the treated unit. Because of the small sample sizes often encountered in practice, Abadie, Diamond, and Hainmueller (2010) propose an in-space placebo test for statistical inference. In addition, Abadie, Diamond, and Hainmueller (2015) recommend an in-time placebo test and a leave-one-out (LOO) robustness test.

While the SCM can be implemented by the excellent command `synth` (Abadie, Diamond, and Hainmueller 2011), it is still difficult for applied researchers to implement placebo tests, even using `synth_runner` (Galiani and Quistorff 2017). The command `allsynth` (Wiltshire 2022) focuses on the bias-corrected version of the SCM and the case with many treated units. However, `allsynth` conducts only the in-space placebo test with p-values based on the mean squared prediction error (MSPE), but no pointwise p-values based on the distribution of placebo effects are provided. Another recent command, `scul` (synthetic control using lasso; see Greathouse [2022]), provides both in-space and in-time placebo tests, but its algorithm uses lasso, ridge, or elastic net to construct counterfactuals. In a sense, `scul` is closer to the regression control method (also known as a panel-data approach to program evaluation; see Hsiao, Ching, and Wan [2012]) and its Stata implementation `rcm` (Yan and Chen 2022) than to the classic SCM.

As a wrapper program for `synth`, our proposed `synth2` command calls on `synth` for implementing the underlying SCM algorithm (Abadie, Diamond, and Hainmueller 2011). Nevertheless, `synth2` provides many convenient functionalities for users that were mostly unavailable in Stata until now. First, `synth2` automates the in-space placebo test for the SCM (Abadie, Diamond, and Hainmueller 2010), which was previously available only in `synth_runner` and `allsynth` in a limited way. Second, `synth2` implements the in-time placebo test for the SCM (Abadie, Diamond, and Hainmueller 2015) for the first time in Stata. Third, `synth2` conducts the LOO robustness test (Abadie, Diamond, and Hainmueller 2015), which is also new in Stata. Finally, `synth2` produces various figures for visualization, many of which were previously unavailable in Stata. These include figures to visualize covariate or unit weights, covariate balance, actual or predicted outcomes, treatment effects, placebo tests, ratio of posttreatment MSPE to pretreatment MSPE (denoted as "post/pre MSPE ratio" for short), pointwise p-values (two-sided, right-sided, and left-sided), and the LOO robustness test.

The rest of the article is organized as follows. Section 2 summarizes the methodology of the SCM. Section 3 and Section 4 discuss placebo tests and the LOO robustness test, respectively. Section 5 presents the command `synth2`. Section 6 illustrates its use by revisiting the classic example of California's tobacco control program (Abadie, Diamond, and Hainmueller 2010). Section 7 concludes.

2 SCM

The exposition of the SCM in this section largely follows Abadie, Diamond, and Hainmueller (2010) and is provided for completeness. Suppose there are $N+1$ cross-sectional units indexed by $i = 1, \ldots, N+1$ and observed over periods $t = 1, \ldots, T_0$ (preintervention) and $t = T_0 + 1, \ldots, T$ (postintervention). To simplify notation, assume the first unit with $i = 1$ to be the treated unit (exposed to the intervention), while the other units with $i = 2, \ldots, N+1$ are control units (not exposed to the intervention) that form the "donor pool". Let y_{it}^1 and y_{it}^0 be the outcomes of unit i in period t with and without intervention, respectively; the observed outcome y_{it} can then be expressed as

$$\begin{aligned} y_{it} &= y_{it}^1 D_{it} + y_{it}^0(1 - D_{it}) \\ &= y_{it}^0 + \Delta_{it} D_{it} \end{aligned}$$

where D_{it} is a treatment indicator such that $D_{it} = 1$ if unit i is treated in period t and $D_{it} = 0$ otherwise. $\Delta_{it} = y_{it}^1 - y_{it}^0$ denotes the treatment effect for unit i in period t. The goal is to estimate $(\Delta_{1T_0+1}, \ldots, \Delta_{1T})$, which is equivalent to estimating $\left(y_{1T_0+1}^0, \ldots, y_{1T}^0\right)$, because $\left(y_{1T_0+1}^1, \ldots, y_{1T}^1\right)$ are observed. Suppose that y_{it}^0 is generated by a factor model

$$y_{it}^0 = \delta_t + \boldsymbol{\theta}_t' \boldsymbol{z}_i + \boldsymbol{\lambda}_t' \boldsymbol{\mu}_i + \varepsilon_{it}$$

where δ_t is a time fixed effect (that is, an unknown common factor with constant factor loadings across units), $\boldsymbol{z}_i$ is a $(K \times 1)$ vector of observed covariates, $\boldsymbol{\theta}_t$ is a $(K \times 1)$ vector of unknown coefficients, $\boldsymbol{\lambda}_t$ is a vector of unobserved common factors, $\boldsymbol{\mu}_i$ is a vector of unknown factor loadings, and ε_{it} is an idiosyncratic shock with a zero mean. The SCM seeks to approximate the unknown y_{1t}^0 $(t = T_0 + 1, \ldots, T)$ by a weighted average of donor units, and the treatment effects are estimated accordingly by

$$\widehat{\Delta}_{1t} = y_{1t} - \widehat{y}_{1t}^0 = y_{1t} - \sum_{i=2}^{N+1} w_i y_{it} \quad (t = T_0 + 1, \ldots, T) \tag{1}$$

Let $\mathbf{w} = (w_2, \ldots, w_{N+1})'$ be a $(N \times 1)$ vector of weights (a potential synthetic control) such that $0 \leq w_i \leq 1$ for $i = 2, \ldots, N+1$ and $\sum_{i=2}^{N+1} w_i = 1$. The SCM selects the optimal $\mathbf{w}$ so that the pretreatment characteristics of the synthetic control are most similar to those of the treated unit. Let $\mathbf{x}_1$ be the $(K \times 1)$ vector containing the pretreatment covariates (predictors) of the treated unit, which may include pretreatment values of outcome, and let $\mathbf{X}_0$ be the $(K \times N)$ matrix containing the pretreatment covariates of the N control units. Moreover, let $\mathbf{V}$ be a $(K \times K)$ diagonal matrix with nonnegative elements on its diagonal that contains covariate weights measuring the importance of each covariate in predicting the outcome. We use the notation $\|\mathbf{x}\|_{\mathbf{V}} \equiv \sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}}$ as a distance measure indexed by $\mathbf{V}$. In particular, if $\mathbf{V}$ is the identity matrix, then it reduces to the usual Euclidean norm $\|\mathbf{x}\| \equiv \sqrt{\mathbf{x}'\mathbf{x}}$. The optimal synthetic control $\mathbf{w}^*(\mathbf{V})$ is obtained by solving the following minimization problem:

$$\mathbf{w}^*(\mathbf{V}) = \arg\min_{\mathbf{w}} \|\mathbf{x}_1 - \mathbf{X}_0 \mathbf{w}\|_{\mathbf{V}}$$

Let $\mathbf{z}_1$ be the $(T_0 \times 1)$ vector of pretreatment outcomes for the treated unit, and let $\mathbf{Z}_0$ be the $(T_0 \times N)$ matrix of pretreatment outcomes for the N control units. Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010) present a data-driven procedure to choose the optimal $\mathbf{V}^*$ that minimizes the MSPE of the outcome variable for the pretreatment period:

$$\mathbf{V}^* = \arg\min_{\mathbf{V}} \|\mathbf{z}_1 - \mathbf{Z}_0 \mathbf{w}^*(\mathbf{V})\|$$

Given the $\mathbf{V}^*$ containing optimal covariate weights, the optimal unit weights $\mathbf{w}^* = \mathbf{w}^*(\mathbf{V}^*)$ can be computed. Thus, we can use the optimal unit weights $\mathbf{w}^*$ to estimate the counterfactual outcome $\widehat{y}^0_{1t}$ and the treatment effect $\widehat{\Delta}_{1t} = y_{1t} - \widehat{y}^0_{1t}$ over the post-treatment period according to (1).

3 Placebo tests

The conventional statistical inference for the SCM relies on placebo tests, which typically come in two forms, that is, the in-space placebo test (Abadie, Diamond, and Hainmueller 2010) and the in-time placebo test (Abadie, Diamond, and Hainmueller 2015). The `synth2` command implements both in-space and in-time placebo tests. The exposition below draws heavily on the above two articles by Abadie, Diamond, and Hainmueller.

3.1 In-space placebo test

The idea of the in-space placebo test is akin to the classic framework for permutation tests, where the distribution of a test statistic is computed under random permutations of the sample units' assignments to the treated and untreated groups. In other words, the in-space placebo test uses "fake treatment units" for statistical inference. Specifically, it compares the estimated treatment effects on the treated unit with a distribution of placebo effects obtained by iteratively assigning the treatment to donor units and estimating placebo effects in each iteration. As a technical detail, we may require the fake treatment units to have a pretreatment MSPE not too much larger (say, 5 or 20 times more) than that of the treated unit. Simply put, fake treatment units with a poor pretreatment fit are dropped because they contain little useful information. The pointwise in-space placebo test considers the null hypothesis

$$H_0 : \Delta_{1t} = 0$$

where Δ_{1t} is the treatment effect for the first unit in period $t = T_0 + 1, \ldots, T$. The treatment effect is considered significant if the estimated treatment effect is "unusually extreme" (either unusually large, small, or large in absolute value) relative to the distribution of placebo effects. Otherwise, the null hypothesis of "no treatment effect" is accepted. Depending on how one measures unusual extremeness, the `synth2` command computes a right-sided p-value (for "unusually large"), a left-sided p-value (for "unusually small", that is, a negative number with a large absolute value), and a two-sided p-value (for "unusually large in absolute value") for each posttreatment period. Specifically, there are three ways to formulate the alternative hypothesis. The first way corresponds to the usual two-tail test:

$$H_1 : \Delta_{1t} \neq 0$$

For this alternative hypothesis, the treatment effect is considered significant if it is unusually large in absolute value relative to the distribution of placebo effects. In particular, one should use the two-sided p-value defined as the frequency that the absolute values of the placebo effects are greater than or equal to the absolute value of the estimated treatment effect

$$\text{two-sided } p\text{-value}(t) = \frac{1}{N+1}\sum_{i=1}^{N+1} \mathbf{1}\left(\left|\widehat{\Delta}_{it}\right| \geq \left|\widehat{\Delta}_{1t}\right|\right), \quad t = T_0+1,\ldots,T$$

where $\widehat{\Delta}_{it}$ is the estimated treatment (placebo) effect for unit i in period t (that is, $\widehat{\Delta}_{1t}$ is the treatment effect, whereas $\widehat{\Delta}_{it}$ is the placebo effect for unit $i \neq 1$); and $\mathbf{1}(\cdot)$ is the indicator function, which equals 1 if the expression inside is true and 0 otherwise. The second way to formulate the alternative hypothesis corresponds to the right-tail test, where the rejection region locates toward the right tail of the distribution:

$$H_2 : \Delta_{1t} > 0$$

Here the possibility of $\Delta_{1t} < 0$ is ruled out a priori, perhaps on a theoretical ground or because the estimated treatment effect is positive and very large. In this case, the treatment effect is considered significant if the estimated treatment effect is unusually large relative to the distribution of placebo effects. Specifically, one should use the right-sided p-value defined as the frequency that the placebo effects are greater than or equal to the estimated treatment effect:

$$\text{right-sided } p\text{-value}(t) = \frac{1}{N+1}\sum_{i=1}^{N+1} \mathbf{1}\left(\widehat{\Delta}_{it} \geq \widehat{\Delta}_{1t}\right), \quad t = T_0+1,\ldots,T$$

The third way to formulate the alternative hypothesis corresponds to the left-tail test, where the rejection region locates toward the left tail of the distribution:

$$H_3 : \Delta_{1t} < 0$$

Now the possibility of $\Delta_{1t} > 0$ is ruled out beforehand, perhaps for a theoretical reason, or because the estimated treatment effect is negative and very small. In this case, the treatment effect is considered significant if the estimated treatment effect is unusually small relative to the distribution of placebo effects. Specifically, one should use the left-sided p-values defined as the frequency that the placebo effects are smaller than or equal to the estimated treatment effect:

$$\text{left-sided } p\text{-value}(t) = \frac{1}{N+1}\sum_{i=1}^{N+1} \mathbf{1}\left(\widehat{\Delta}_{it} \leq \widehat{\Delta}_{1t}\right), \quad t = T_0+1,\ldots,T$$

In general, one-sided p-values (right-sided or left-sided) provide more power than two-sided p-values. For example, if the estimated treatment effects are all positive,

then one may rule out the possibility of negative treatment effects. Consequently, one could use right-sided p-values for right-sided tests for best results. On the contrary, if the estimated treatment effects are all negative, then left-sided p-values for left-sided tests are recommended for the same reason. Moreover, if the estimated treatment effects fluctuate between the positive and negative territories, then one may choose the smallest p-value out of the three p-values for each posttreatment period.

The above p-values measure pointwise significance of the treatment effects. As an overall measure of the significance of treatment effects over the entire posttreatment period, we can compare the post/pre MSPE ratio for the treated unit with a placebo distribution of this ratio obtained by the above in-space placebo test. Intuitively, if the post/pre MSPE ratio for the treated unit is unusually large relative to the placebo distribution of this ratio, then we are more confident that the overall treatment effects are significant. Specifically, the probability (that is, p-value) of obtaining a post/pre MSPE ratio as large as that of the treated unit is calculated as

$$\text{MSPE-based } p\text{-value} = \frac{1}{N+1}\sum_{i=1}^{N+1} \mathbf{1}\left(\frac{\text{MSPE}_{i,\text{post}}}{\text{MSPE}_{i,\text{pre}}} \geq \frac{\text{MSPE}_{1,\text{post}}}{\text{MSPE}_{1,\text{pre}}}\right)$$

where $\text{MSPE}_{i,\text{post}}$ and $\text{MSPE}_{i,\text{pre}}$ are posttreatment MSPE and pretreatment MSPE for unit i, respectively. For example, if the post/pre MSPE ratio for the treated unit is larger than all control units, then the corresponding p-value is $1/(N+1)$.

3.2 In-time placebo test

The in-time placebo test makes use of a fake treatment time before the treatment actually starts, which is also known as "backdating". Specifically, a fake treatment time in the pretreatment period is chosen, say, $\widetilde{T}_0 < T_0 + 1$ (the actual treatment starts in period $T_0 + 1$). We then assign the treatment to unit 1 from period $\widetilde{T}_0$ on, where no treatment actually occurred during the period $[\widetilde{T}_0, T_0]$.

The intuition is that if the estimated placebo effects during the period $[\widetilde{T}_0, T_0]$ turn out to be "significant" or "large" in some sense, then it erodes our confidence in the significance of the actual treatment effects. Note that no p-value is computed for the in-time placebo test, and one typically uses a graph to present the results from the in-time placebo test. In addition, a researcher can choose multiple fake treatment times and conduct in-time placebo tests for each fake treatment time separately.

4 Robustness test

The `synth2` command also implements the LOO robustness test proposed by Abadie, Diamond, and Hainmueller (2015). As a weighted average of donor units, the optimal synthetic control typically is sparse in that most control units receive a zero weight. Therefore, one may be concerned that the estimated treatment effects may be disproportionally driven by just one control unit with a nonzero weight.

The LOO robustness test reestimates the original SCM by omitting one of the original selected donors in each iteration. Intuitively, the LOO analysis evaluates to what extent results are driven by any particular control unit, although excluding a non-zero-weighted unit sacrifices some goodness of fit. If the outcomes and treatment effects of LOO synthetic controls are similar to those of synthetic controls with all control units, then the estimated results are considered robust.

5 The synth2 command

5.1 Syntax

The syntax for `synth2` is similar to `synth` but augmented with additional options to implement placebo and robustness tests:

`synth2` *depvar indepvars*`,` `trunit(`#`)` `trperiod(`#`)` [`counit(`*numlist*`)`
`preperiod(`*numlist*`)` `postperiod(`*numlist*`)` `xperiod(`*numlist*`)`
`mspeperiod(`*numlist*`)` `nested allopt` `customV(`*numlist*`)` `margin(`*real*`)`
`maxiter(`#`)` `sigf(`#`)` `bound(`#`)`
`placebo(`[[`unit`|`unit(`*numlist*`)`] `period(`*numlist*`)` `cutoff(`$\#_c$`)` `show(`$\#_s$`)`]`)`
`loo` `frame(`*framename*`)` `nofigure` `savegraph(`*prefix*[`, asis replace`]`)`]

`xtset` *panelvar timevar* must be used to declare a balanced panel dataset in the usual long form; see [XT] **xtset**.

depvar and *indepvars* must be numeric variables, and abbreviations are not allowed. *indepvars* may include lagged values of *depvar* specified as *depvar*`(`*period*`)`.

5.2 Options

Some options below are identical to those of `synth`, and they share the same option names. On the other hand, a different option name signifies a unique option specific to `synth2`. Note that some important options are explained below for completeness, despite being identical with those of `synth`; otherwise, the reader is referred to `synth`.

5.2.1 Required settings

`trunit(`#`)` specifies the unit number of the treated unit (that is, the unit affected by the intervention) as given in the panel variable specified in `xtset` *panelvar*. Note that only one unit number can be specified.

`trperiod(`#`)` specifies the time period when the intervention occurred. The time period refers to the time variable specified in `xtset` *timevar* and must be an integer (see examples below). Note that only one time period can be specified.

5.2.2 Model

`counit(`*numlist*`)` specifies a list of unit numbers for the control units as *numlist* given in the panel variable specified in `xtset` *panelvar*. The list of control units specified constitutes what is known as the "donor pool". The donor pool defaults to all available units other than the treated unit.

`preperiod(`*numlist*`)` specifies a list of pretreatment periods as *numlist* given in the time variable specified in `xtset` *timevar*. `preperiod()` defaults to the entire preintervention period, which ranges from the earliest time period available in the time variable to the period immediately prior to the intervention.

`postperiod(`*numlist*`)` specifies a list of posttreatment periods (when and after the intervention occurred) as *numlist* given in the time variable specified in `xtset` *timevar*. `postperiod()` defaults to the entire postintervention period, which ranges from the time period when the intervention occurred to the latest time period available in the time variable.

`xperiod(`*numlist*`)` specifies a list of periods as *numlist* given in the time variable specified in `xtset` *timevar*, over which the covariates specified in *indepvars* are averaged.

`mspeperiod(`*numlist*`)` specifies a list of pretreatment periods as *numlist* given in the time variable specified in `xtset` *timevar*, over which the MSPE should be minimized.

`nested`, if specified, will have `synth2` embark on a fully nested optimization procedure, which achieves better accuracy than the default algorithm at the expense of additional computing time. For details, see `synth`.

`allopt` provides a robustness check by running the nested optimization three times using three different starting points and returns the best result. If `nested` is specified, the user can also specify `allopt` if he or she is willing to trade off even more computing time to gain fully robust results. For details, see `synth`.

`customV(`*numlist*`)` specifies a list of custom V-weights as *numlist* appearing in the same order as the covariates listed in *indepvars* to replace the data-driven V-weights. For details, see `synth`.

5.2.3 Optimization

`synth2` uses `synth`'s constrained quadratic optimization routine. The options `margin()`, `maxiter()`, `sigf()`, and `bound()` are identical to those of the `synth` command, and the reader is referred to `synth`.

5.2.4 Placebo tests

`placebo(`[[`unit` | `unit(`*numlist*`)`] `period(`*numlist*`)` `cutoff(`$\#_c$`)` `show(`$\#_s$`)`]`)` specifies the types of placebo tests to be performed; otherwise, no placebo test will be implemented.

`unit` or `unit(`*numlist*`)` specifies the in-space placebo test using fake treatment units in the donor pool, where `unit` uses all fake treatment units and `unit(`*numlist*`)` uses a list of fake treatment units specified by *numlist*. These two options iteratively reassign the treatment to control units where no intervention actually occurred and calculate the p-values of the treatment effects. Note that only one of `unit` and `unit()` can be specified.

`period(`*numlist*`)` specifies the in-time placebo test using fake treatment times (more than one fake treatment time can be specified). This option reassigns the treatment to time periods previous to the intervention, when no treatment actually occurred.

`cutoff(`$\#_c$`)` specifies a cutoff threshold that discards fake treatment units with pretreatment MSPE $\#_c$ times larger than that of the treated unit, where $\#_c$ must be a real number greater than or equal to 1. This option applies only when `unit` or `unit()` is specified. By default, no fake treatment units are discarded.

`show(`$\#_s$`)` specifies the number of units to show in the post/pre MSPE graph, which corresponds to units with the largest $\#_s$ ratios of posttreatment MSPE to pretreatment MSPE. This option applies only when `unit` or `unit()` is specified. The default is to show post/pre MSPE ratios for all units.

5.2.5 Robustness test

`loo` specifies the LOO robustness test that excludes one control unit with a nonzero weight at a time. `synth2` iteratively refits the model, omitting one unit in each iteration that receives a positive weight. By excluding a unit receiving a positive weight, goodness of fit is sacrificed, but this sensitivity check can evaluate the extent to which results are driven by any particular control unit.

5.2.6 Reporting

`frame(`*framename*`)` creates a Stata frame that stores generated variables in the wide form, including counterfactual predictions, treatment effects, and results from placebo tests if implemented. The frame named *framename* is replaced if it already exists or created if it does not.

`nofigure` specifies not to display figures. The default is to display all figures from the estimation results, placebo tests, and robustness test if available.

`savegraph(`*prefix*`[ , asis replace ])` automatically and iteratively calls `graph save` to save all produced graphs to the current path, where *prefix* specifies the prefix added to *_graphname* to form a filename; that is, the graph named *graphname* is stored as *prefix_graphname*`.gph`. `asis` and `replace` are options passed to `graph save`; for details, see [G-2] **graph save**. Note that this option applies only when `nofigure` is not specified.

5.3 Stored results

`synth2` stores the following in `e()`:

Scalars

`e(N)`	number of observations
`e(T0)`	number of pretreatment periods
`e(T1)`	number of posttreatment periods
`e(K)`	number of covariates
`e(rmse)`	root mean squared error of the model fit in the pretreatment period
`e(r2)`	R^2 of the model fit over the posttreatment period
`e(att)`	average treatment effect

Macros

`e(panelvar)`	name of the panel variable
`e(timevar)`	name of the time variable
`e(varlist)`	names of the dependent variable and independent variables
`e(depvar)`	name of dependent variable
`e(indepvars)`	names of independent variables (covariates)
`e(unit_all)`	all units
`e(unit_tr)`	treatment unit
`e(unit_ctrl)`	control units
`e(time_all)`	entire periods
`e(time_tr)`	treatment period
`e(time_pre)`	pretreatment periods
`e(time_post)`	posttreatment periods
`e(frame)`	name of Stata frame storing generated variables
`e(graph)`	names of all produced graphs

Matrices

`e(V_wt)`	diagonal matrix **V** containing the optimal covariate weights in the diagonal
`e(U_wt)`	vector **w** that contains the optimal unit weights
`e(bal)`	matrix containing sample averages for the treated unit, synthetic control unit, and control units
`e(mspe)`	matrix containing pretreatment MSPE, posttreatment MSPE, ratios of posttreatment MSPE to pretreatment MSPE, and ratios of pretreatment MSPE of control units to that of the treated unit
`e(pval)`	matrix containing estimated treatment effects and p-values from placebo tests using fake treatment units

6 Examples

6.1 Example 1: Replicate Abadie, Diamond, and Hainmueller (2010)

To demonstrate the use of `synth2`, we replicate the classic example about the effect of California's tobacco control program (Proposition 99) on cigarette sales (Abadie, Diamond, and Hainmueller 2010). `smoking.dta` attached to the `synth2` command includes the following variables for 39 U.S. States from 1970 to 2000: the outcome variable `cigsale` (cigarette sale per capita in packs) and covariates `lnincome` (logged per-capita state personal income), `age15to24` (percentage of the population aged 15–24), `retprice` (annual state-level values of average retail price of cigarettes), and `beer` (per-capita beer consumption).

After loading `smoking.dta`, we declare it as a panel dataset:

```
. use smoking
(Tobacco Sales in 39 US States)
. xtset state year
Panel variable: state (strongly balanced)
 Time variable: year, 1970 to 2000
         Delta: 1 unit
```

Next, we use the command `label list` to find the unit number for the treated unit, California:

```
. label list
state:
           1 Alabama
           2 Arkansas
           3 California
           4 Colorado
           5 Connecticut
           6 Delaware
           7 Georgia
           8 Idaho
           9 Illinois
          10 Indiana
          11 Iowa
          12 Kansas
          13 Kentucky
          14 Louisiana
          15 Maine
          16 Minnesota
          17 Mississippi
          18 Missouri
          19 Montana
          20 Nebraska
          21 Nevada
          22 New Hampshire
          23 New Mexico
          24 North Carolina
          25 North Dakota
          26 Ohio
          27 Oklahoma
          28 Pennsylvania
          29 Rhode Island
          30 South Carolina
          31 South Dakota
          32 Tennessee
          33 Texas
          34 Utah
          35 Vermont
          36 Virginia
          37 West Virginia
          38 Wisconsin
          39 Wyoming
```

The results show that the unit number for California is 3. Hence, we use the option `trunit(3)` to specify California as the treated unit.

To specify the treatment period, we use the option `trperiod(1989)` because California's tobacco control legislation was passed in November 1988 and became effective in January 1989. Following Abadie, Diamond, and Hainmueller (2010), we use the option `xperiod(1980(1)1988)` to average the covariates over the 1980–1988 periods[1] and include covariates `cigsale(1988)`, `cigsale(1980)`, and `cigsale(1975)`, which are the values of `cigsale` in 1988, 1980, and 1975, respectively. Moreover, we use the options `nested` and `allopt` to produce the most accurate results at the expense of extra computing time.

After collecting all the above information, we use the `synth2` command to replicate the results of Abadie, Diamond, and Hainmueller (2010):

```
. synth2 cigsale lnincome age15to24 retprice beer cigsale(1988) cigsale(1980)
> cigsale(1975), trunit(3) trperiod(1989) xperiod(1980(1)1988) nested allopt
Fitting results in the pretreatment periods:
```

Treated Unit	: California	Treatment Time	: 1989
Number of Control Units	= 38	Root Mean Squared Error	= 1.75567
Number of Covariates	= 7	R-squared	= 0.97434

Covariate balance in the pretreatment periods:

Covariate	V.weight	Treated	Synthetic Control		Average Control	
			Value	Bias	Value	Bias
lnincome	0.0000	10.0766	9.8588	-2.16%	9.8292	-2.45%
age15to24	0.5459	0.1735	0.1735	-0.01%	0.1725	-0.59%
retprice	0.0174	89.4222	89.4108	-0.01%	87.2661	-2.41%
beer	0.0031	24.2800	24.2278	-0.21%	23.6553	-2.57%
cigsale(1988)	0.0049	90.1000	91.6677	1.74%	113.8237	26.33%
cigsale(1980)	0.0066	120.2000	120.5017	0.25%	138.0895	14.88%
cigsale(1975)	0.4221	127.1000	127.1112	0.01%	136.9316	7.74%

```
Note: "V.weight" is the optimal covariate weight in the diagonal of V matrix.
      "Synthetic Control" is the weighted average of donor units with optimal
      weights.
      "Average Control" is the simple average of all control units with equal
      weights.
```

1. Because the `beer` variable has no observation before 1984, this is equivalent to averaging over the 1984–1988 period for the `beer` variable.

```
Optimal Unit Weights:
```

Unit	U.weight
Utah	0.3340
Nevada	0.2350
Montana	0.2020
Colorado	0.1610
Connecticut	0.0680

```
Note: The unit Alabama Arkansas Delaware Georgia Idaho Illinois Indiana Iowa
      Kansas Kentucky Louisiana Maine Minnesota Mississippi Missouri Nebraska
      NewHampshire NewMexico NorthCarolina NorthDakota Ohio Oklahoma
      Pennsylvania RhodeIsland SouthCarolina SouthDakota Tennessee Texas
      Vermont Virginia WestVirginia Wisconsin Wyoming in the donor pool get a
      weight of 0.
Prediction results in the posttreatment periods:
```

Time	Actual Outcome	Synthetic Outcome	Treatment Effect
1989	82.4000	89.9945	-7.5945
1990	77.8000	87.5039	-9.7039
1991	68.7000	82.1751	-13.4751
1992	67.5000	81.6075	-14.1075
1993	63.4000	81.1897	-17.7897
1994	58.6000	80.7295	-22.1295
1995	56.4000	78.5023	-22.1023
1996	54.5000	77.4827	-22.9827
1997	53.8000	77.7123	-23.9123
1998	52.3000	74.3976	-22.0976
1999	47.2000	73.5711	-26.3711
2000	41.6000	67.3550	-25.7550
Mean	60.3500	79.3518	-19.0018

```
Note: The average treatment effect over the posttreatment period is -19.0018.
Finished.
```

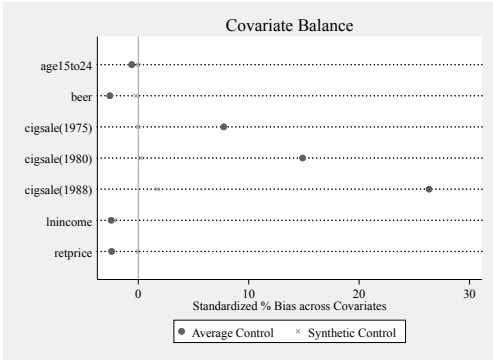
The above results show an excellent pretreatment fit, where the R^2 reaches 0.97434. The optimal covariate weights (reported as `V.weight` above) indicate that `age15to24` and `cigsale(1975)` receive much larger weights than other covariates.

In terms of replicating the pretreatment characteristics of the treated unit, the "synthetic control" (a weighted average of donor units with optimal weights) achieves a great covariate balance such that the largest covariate difference in percentage in absolute value between actual and synthetic California is only 2.16% for `lnincome`, which is reported as "bias" in the covariate balance table and computed as $(9.8588 - 10.0766)/10.0766$. In contrast, if an "average control" (a simple average of all control units with equal weights) is used, the largest covariate difference reaches 26.33% for `cigsale(1988)`.

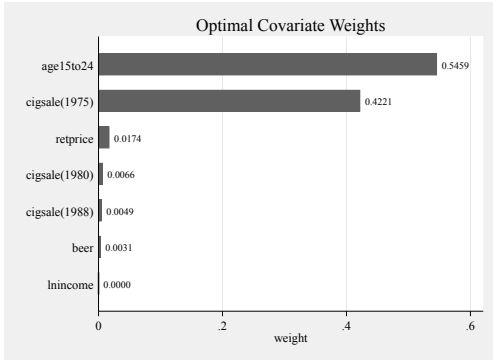
The optimal unit weights (reported as `U.weight` above) reveal that the synthetic control for California consists of a convex combination of Utah, Nevada, Montana, Colorado, and Connecticut, whereas all other control units receive zero weights. The actual outcomes, predicted outcomes, and treatment effects are also reported for each posttreatment period.

In the meantime, the above `synth2` command produces five graphs collected in figure 1.[2] Figure 1(a) contrasts the covariate balance between the synthetic control and the average control, where the gray vertical line represents the treated unit. Figure 1(b) presents the optimal covariate weights (the diagonal elements of matrix $\mathbf{V}^*$) in a horizontal bar graph. Similarly, figure 1(c) graphs the optimal unit weights (the weight vector $\mathbf{w}^*$). Figure 1(d) depicts the actual and predicted outcomes, also known as the "gap graph". Finally, figure 1(e) provides a visualization of the estimated treatment effects.

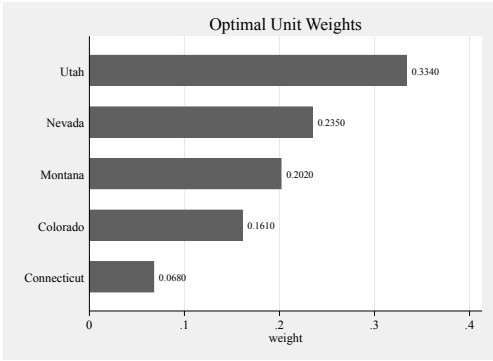
2. To save space, we combine these graphs into one chart. Commands for retrieving this chart and other charts containing multiple graphs are provided in the help file and the `example.do` file.



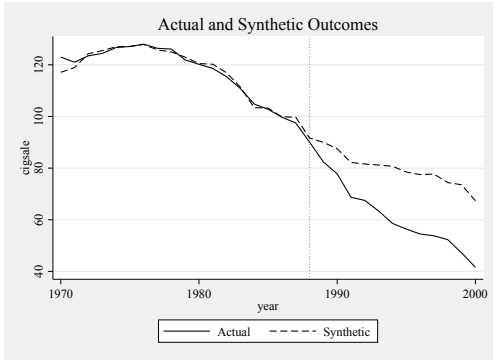
(a) Covariate balance



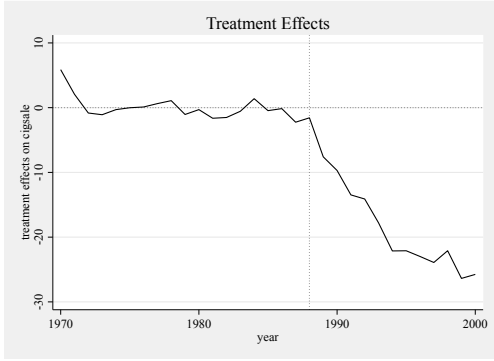
(b) Optimal covariate weights (diagonal of V^*)



(c) Optimal unit weights (W^*)



(d) Actual and predicted paths



(e) Treatment effects

Figure 1. Graphs for California's tobacco control program in example 1

6.2 Example 2: In-space placebo test

In this example, we implement the in-space placebo test. The option `placebo(unit cutoff(2))` is added to request the in-space placebo test using all fake treatment units but exclude those units with pretreatment MSPEs two times larger than that of the treated unit. Note that one can also replace `unit` with `unit()` in this option to specify candidate control units as fake treatment units. We drop the `allopt` option to save time but still keep the `nested` option for accuracy. In addition, we change the default option `sigf(7)` (7 significant figures) to `sigf(6)` to ensure convergence. Implementing the following command may be time consuming, but it is certainly worth the wait.

```
. synth2 cigsale lnincome age15to24 retprice beer cigsale(1988) cigsale(1980)
> cigsale(1975), trunit(3) trperiod(1989) xperiod(1980(1)1988) nested
> placebo(unit cutoff(2)) sigf(6)
Fitting results in the pretreatment periods:
```

Treated Unit	: California	Treatment Time	: 1989
Number of Control Units	= 38	Root Mean Squared Error	= 1.77391
Number of Covariates	= 7	R-squared	= 0.97329

Covariate balance in the pretreatment periods:

Covariate	V.weight	Treated	Synthetic Control		Average Control	
			Value	Bias	Value	Bias
lnincome	0.0009	10.0766	9.8528	-2.22%	9.8292	-2.45%
age15to24	0.0153	0.1735	0.1735	-0.04%	0.1725	-0.59%
retprice	0.0910	89.4222	89.2616	-0.18%	87.2661	-2.41%
beer	0.0250	24.2800	24.1186	-0.66%	23.6553	-2.57%
cigsale(1988)	0.1015	90.1000	91.5412	1.60%	113.8237	26.33%
cigsale(1980)	0.0583	120.2000	120.3496	0.12%	138.0895	14.88%
cigsale(1975)	0.7080	127.1000	126.7802	-0.25%	136.9316	7.74%

```
Note: "V.weight" is the optimal covariate weight in the diagonal of V matrix.
      "Synthetic Control" is the weighted average of donor units with optimal
      weights.
      "Average Control" is the simple average of all control units with equal
      weights.
```

Optimal Unit Weights:

Unit	U.weight
Utah	0.3320
Nevada	0.2300
Montana	0.1880
Colorado	0.1790
Connecticut	0.0700

```
Note: The unit Alabama Arkansas Delaware Georgia Idaho Illinois Indiana Iowa
      Kansas Kentucky Louisiana Maine Minnesota Mississippi Missouri Nebraska
      NewHampshire NewMexico NorthCarolina NorthDakota Ohio Oklahoma
      Pennsylvania RhodeIsland SouthCarolina SouthDakota Tennessee Texas
      Vermont Virginia WestVirginia Wisconsin Wyoming in the donor pool get a
      weight of 0.
```

Prediction results in the posttreatment periods:

Time	Actual Outcome	Synthetic Outcome	Treatment Effect
1989	82.4000	89.7838	-7.3838
1990	77.8000	87.2810	-9.4810
1991	68.7000	82.1270	-13.4270
1992	67.5000	81.4915	-13.9915
1993	63.4000	81.0858	-17.6858
1994	58.6000	80.6141	-22.0141
1995	56.4000	78.3226	-21.9226
1996	54.5000	77.3053	-22.8053
1997	53.8000	77.4909	-23.6909
1998	52.3000	74.1662	-21.8662
1999	47.2000	73.3870	-26.1870
2000	41.6000	67.2074	-25.6074
Mean	60.3500	79.1885	-18.8385

Note: The average treatment effect over the posttreatment period is -18.8385.

```
Implementing placebo test using fake treatment unit Alabama...Arkansas...
> Colorado...Connecticut...Delaware...Georgia...Idaho...Illinois...Indiana...
> Iowa...Kansas...Kentucky...Louisiana...Maine...Minnesota...Mississippi...
> Missouri...Montana...Nebraska...Nevada...NewHampshire...NewMexico...
> NorthCarolina...NorthDakota...Ohio...Oklahoma...Pennsylvania...RhodeIsland...
> SouthCarolina...SouthDakota...Tennessee...Texas...Utah...Vermont...Virginia...
> WestVirginia...Wisconsin...Wyoming...
```

In-space placebo test results using fake treatment units:

Unit	Pre MSPE	Post MSPE	Post/Pre MSPE	Pre MSPE of Fake Unit/ Pre MSPE of Treated Unit
California	3.1467	391.6195	124.4523	1.0000
Alabama	5.1122	6.8512	1.3402	1.6246
Arkansas	4.5460	26.9649	5.9316	1.4447
Colorado	15.3826	53.9001	3.5040	4.8884
Connecticut	20.8269	12.6260	0.6062	6.6185
Delaware	42.4554	467.4360	11.0100	13.4919
Georgia	1.5158	114.3227	75.4215	0.4817
Idaho	5.5430	40.2997	7.2704	1.7615
Illinois	8.0931	56.8125	7.0199	2.5719
Indiana	14.2226	478.6144	33.6517	4.5198
Iowa	13.7582	28.1012	2.0425	4.3722
Kansas	13.9257	14.4616	1.0385	4.4254
Kentucky	431.9344	1497.8958	3.4679	137.2639
Louisiana	1.9390	99.8782	51.5091	0.6162
Maine	9.3788	143.7055	15.3224	2.9805
Minnesota	15.0327	44.0872	2.9327	4.7772
Mississippi	3.9232	37.2858	9.5039	1.2467
Missouri	1.2576	77.1025	61.3076	0.3997
Montana	5.2862	54.8978	10.3851	1.6799
Nebraska	4.3496	44.6502	10.2655	1.3822
Nevada	40.6733	83.5320	2.0537	12.9255
NewHampshire	3436.5977	134.9018	0.0393	1092.1125
NewMexico	5.0860	67.1420	13.2014	1.6163
NorthCarolina	81.5801	58.8357	0.7212	25.9253
NorthDakota	8.1963	83.7060	10.2126	2.6047
Ohio	5.2247	14.7285	2.8190	1.6604
Oklahoma	4.8431	240.6424	49.6878	1.5391
Pennsylvania	2.8199	7.2463	2.5697	0.8961
RhodeIsland	67.3356	217.6933	3.2330	21.3985
SouthCarolina	2.2061	41.4957	18.8096	0.7011
SouthDakota	7.1493	33.8704	4.7376	2.2720
Tennessee	5.2043	123.3097	23.6937	1.6539
Texas	4.6983	237.2759	50.5020	1.4931
Utah	593.7643	223.2758	0.3760	188.6917
Vermont	16.4634	117.0065	7.1071	5.2319
Virginia	2.7749	162.1030	58.4179	0.8818
WestVirginia	8.6441	226.6917	26.2251	2.7470
Wisconsin	2.7290	83.8542	30.7275	0.8672
Wyoming	83.6674	76.4727	0.9140	26.5886

Note: (1) Using all control units, the probability of obtaining a post/pretreatment MSPE ratio as large as California's is 0.0256.
(2) Excluding control units with pretreatment MSPE 2 times larger than the treated unit, the probability of obtaining a post/pretreatment MSPE ratio as large as California's is 0.0526.
(3) The pointwise p-values below are computed by excluding control units with pretreatment MSPE 2 times larger than the treated unit.
(4) There are total 20 units with pretreatment MSPE 2 times larger than the treated unit, including Colorado Connecticut Delaware Illinois Indiana Iowa Kansas Kentucky Maine Minnesota Nevada NewHampshire NorthCarolina NorthDakota RhodeIsland SouthDakota Utah Vermont WestVirginia Wyoming.

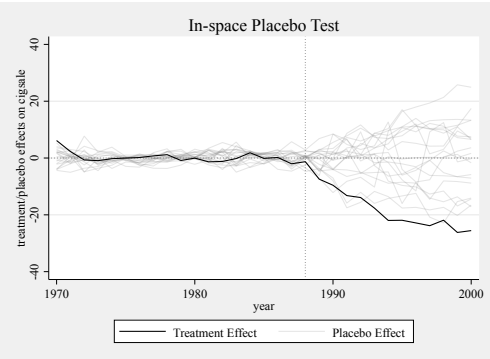
```
In-space placebo test results using fake treatment units (continued, cutoff = 2):
```

Time	Treatment Effect	p-value of Treatment Effect		
		Two-sided	Right-sided	Left-sided
1989	-7.3838	0.0526	1.0000	0.0526
1990	-9.4810	0.1053	0.9474	0.1053
1991	-13.4270	0.1579	0.8947	0.1579
1992	-13.9915	0.1053	0.9474	0.1053
1993	-17.6858	0.0526	1.0000	0.0526
1994	-22.0141	0.0526	1.0000	0.0526
1995	-21.9226	0.0526	1.0000	0.0526
1996	-22.8053	0.0526	1.0000	0.0526
1997	-23.6909	0.0526	1.0000	0.0526
1998	-21.8662	0.0526	1.0000	0.0526
1999	-26.1870	0.0526	1.0000	0.0526
2000	-25.6074	0.0526	1.0000	0.0526

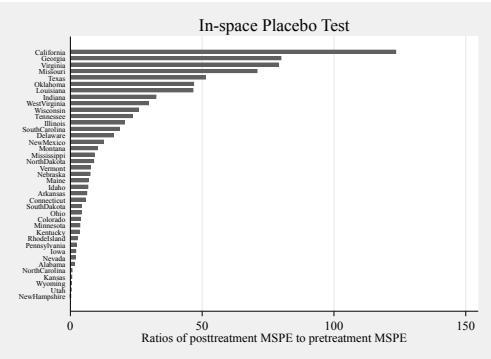
```
Note: (1) The two-sided p-value of the treatment effect for a particular
      period is defined as the frequency that the absolute values of the
      placebo effects are greater than or equal to the absolute value of
      treatment effect.
      (2) The right-sided (left-sided) p-value of the treatment effect for a
      particular period is defined as the frequency that the placebo effects
      are greater (smaller) than or equal to the treatment effect.
      (3) If the estimated treatment effect is positive, then the right-sided
      p-value is recommended; whereas the left-sided p-value is recommended if
      the estimated treatment effect is negative.

Finished.
```

The above results show that California has the largest post/pre MSPE ratio among all 39 states, yielding an overall p-value of $1/39 = 0.0256$, which is significant at the 5% level. Moreover, even if we drop control units with pretreatment MSPEs two times larger than the treated unit, the MSPE-based p-value is still 0.0526. Furthermore, if we look at the pointwise p-values (either two-sided or left-sided p-values), the treatment effects are significant at the 5% level for most posttreatment periods.

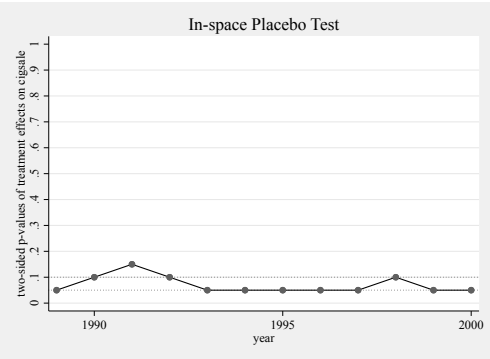
In the meantime, the above `synth2` command produces five graphs collected in figure 2. Figure 2(a) graphs the distribution of placebo effects, against which the estimated treatment effects are compared. Apparently, the estimated treatment effects are all negative and mostly lie at the bottom of the distribution of placebo effects. Figure 2(b) presents the post/pre MSPE ratios in a horizontal bar graph, where the post/pre MSPE ratio for California is clearly the largest. Note that one could use the option `show()` to restrict the number of units to display in this graph; for example, we could specify `placebo(unit cutoff(2) show(10))`. Figures 2(c), 2(d), and 2(e) graph two-sided, right-sided, and left-sided p-values, respectively.



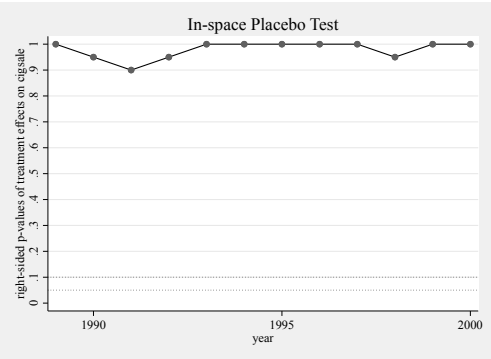
(a) Treatment and placebo effects



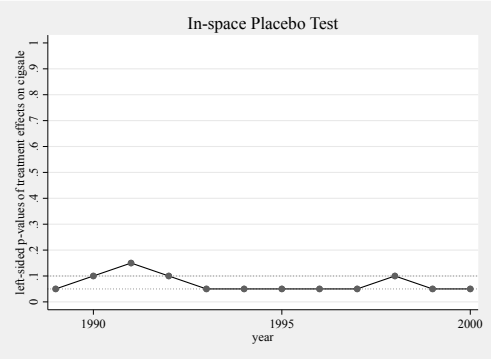
(b) Post/pre MSPE ratios



(c) Two-sided p-values of treatment effects



(d) Right-sided p-values of treatment effects



(e) Left-sided p-values of treatment effects

Figure 2. Graphs for the in-space placebo test in example 2

6.3 Example 3: In-time placebo test

In this example, we implement the in-time placebo test. The `placebo(period(1985))` option specifies the in-time placebo test with 1985 as the fake treatment time, which is four years earlier than the actual treatment time of 1989. In addition, we remove the covariate `cigsale(1988)`, which happened after the posited fake treatment time 1985, and update the option `xperiod(1980(1)1988)` to `xperiod(1980(1)1984)` accordingly. Note that the results are very similar if we replace the covariate `cigsale(1988)` with `cigsale(1984)`, which is unreported to save space.

```
. synth2 cigsale lnincome age15to24 retprice beer cigsale(1980) cigsale(1975),
> trunit(3) trperiod(1989) xperiod(1980(1)1984) nested placebo(period(1985))
Fitting results in the pretreatment periods:
```

Treated Unit	: California		Treatment Time	:	1989
Number of Control Units	=	38	Root Mean Squared Error	=	2.20577
Number of Covariates	=	6	R-squared	=	0.95216

Covariate balance in the pretreatment periods:

Covariate	V.weight	Treated	Synthetic Control		Average Control	
			Value	Bias	Value	Bias
lnincome	0.0000	10.0372	9.8736	-1.63%	9.7892	-2.47%
age15to24	0.0000	0.1815	0.1827	0.65%	0.1814	-0.06%
retprice	0.0728	76.2200	76.2252	0.01%	71.8353	-5.75%
beer	0.0000	25.0000	23.0372	-7.85%	23.6947	-5.22%
cigsale(1980)	0.9028	120.2000	120.1873	-0.01%	138.0895	14.88%
cigsale(1975)	0.0243	127.1000	126.9355	-0.13%	136.9316	7.74%

```
Note: "V.weight" is the optimal covariate weight in the diagonal of V matrix.
      "Synthetic Control" is the weighted average of donor units with optimal
      weights.
      "Average Control" is the simple average of all control units with equal
      weights.
Optimal Unit Weights:
```

Unit	U.weight
Utah	0.3600
Nevada	0.2880
Connecticut	0.1990
Colorado	0.1020
NewMexico	0.0510

```
Note: The unit Alabama Arkansas Delaware Georgia Idaho Illinois Indiana Iowa
      Kansas Kentucky Louisiana Maine Minnesota Mississippi Missouri Montana
      Nebraska NewHampshire NorthCarolina NorthDakota Ohio Oklahoma
      Pennsylvania RhodeIsland SouthCarolina SouthDakota Tennessee Texas
      Vermont Virginia WestVirginia Wisconsin Wyoming in the donor pool get a
      weight of 0.
```

```
Prediction results in the posttreatment periods:
```

Time	Actual Outcome	Synthetic Outcome	Treatment Effect
1989	82.4000	93.1066	-10.7066
1990	77.8000	89.5005	-11.7005
1991	68.7000	82.5426	-13.8426
1992	67.5000	80.7445	-13.2445
1993	63.4000	79.6619	-16.2619
1994	58.6000	78.1954	-19.5954
1995	56.4000	75.6877	-19.2877
1996	54.5000	74.9029	-20.4029
1997	53.8000	74.6013	-20.8013
1998	52.3000	71.2187	-18.9187
1999	47.2000	71.4977	-24.2977
2000	41.6000	65.8920	-24.2920
Mean	60.3500	78.1293	-17.7793

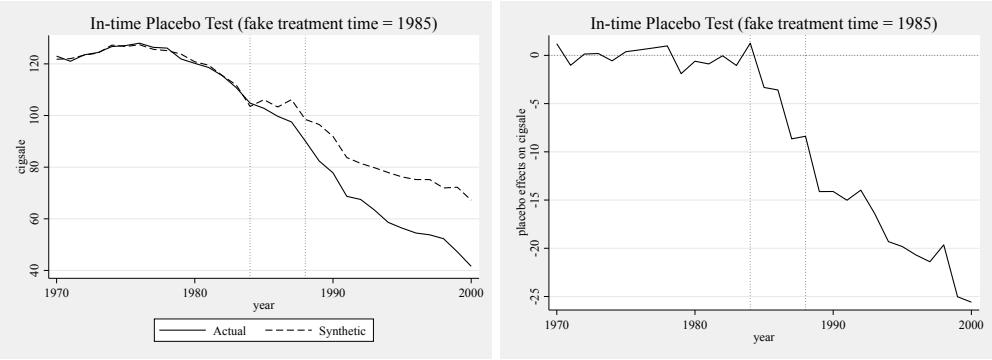
```
Note: The average treatment effect over the posttreatment period is -17.7793.
Implementing placebo test using fake treatment time 1985...
In-time placebo test results using fake treatment time 1985:
```

Time	Actual Outcome	Synthetic Outcome	Treatment Effect
1985	102.8000	106.1262	-3.3262
1986	99.7000	103.2850	-3.5850
1987	97.5000	106.1524	-8.6524
1988	90.1000	98.4873	-8.3873
1989	82.4000	96.5237	-14.1237
1990	77.8000	91.9127	-14.1127
1991	68.7000	83.7156	-15.0156
1992	67.5000	81.4730	-13.9730
1993	63.4000	79.7911	-16.3911
1994	58.6000	77.9078	-19.3078
1995	56.4000	76.2193	-19.8193
1996	54.5000	75.2010	-20.7010
1997	53.8000	75.1958	-21.3958
1998	52.3000	71.9437	-19.6437
1999	47.2000	72.2260	-25.0260
2000	41.6000	67.1861	-25.5861
Mean	69.6437	85.2092	-15.5654

```
Note: The average treatment effect over the posttreatment period is -15.5654.
Finished.
```

The above results report the estimated placebo effects starting from the fake treatment time 1985. More intuitively, the `synth2` command produces two graphs collected in figure 3, where the two dotted vertical lines correspond to the actual and fake treatment times, respectively. Figure 3(a) presents the gap graph with actual and predicted outcomes, pretending that the treatment starts from 1985. There appear to be some noticeable placebo effects during 1985–1988, when there was in fact no treatment. Figure 3(b) provides the corresponding graph for placebo effects, where the "significance" of placebo effects during 1985–1988 appears more obvious. One possible explanation

is that an antismoking movement might have started a few years earlier in California, which culminated in the passage of Proposition 99 in 1988.



(a) Actual and predicted paths (b) Treatment effects

Figure 3. Graphs for the in-time placebo test in example 3

6.4 Example 4: Leave-one-out robustness test

In this example, we implement the LOO robustness test by the option `loo`. Moreover, the option `frame(california)` is specified to create or replace a Stata frame called `california`, which stores generated variables (including predicted outcomes and treatment effects) such that users may find them useful later on (for example, to draw their own figures). In addition, the option `savegraph(california, replace)` is added to save all produced graphs to the current path, where the graph named *graphname* is stored as `california_`*graphname*`.gph`.

```
. synth2 cigsale lnincome age15to24 retprice beer cigsale(1988) cigsale(1980)
> cigsale(1975), trunit(3) trperiod(1989) xperiod(1980(1)1988) nested loo
> frame(california) savegraph(california, replace)
Fitting results in the pretreatment periods:
```

Treated Unit	: California	Treatment Time	: 1989
Number of Control Units	= 38	Root Mean Squared Error	= 1.77768
Number of Covariates	= 7	R-squared	= 0.97423

Covariate balance in the pretreatment periods:

Covariate	V.weight	Treated	Synthetic Control		Average Control	
			Value	Bias	Value	Bias
lnincome	0.0000	10.0766	9.8509	-2.24%	9.8292	-2.45%
age15to24	0.0001	0.1735	0.1735	-0.00%	0.1725	-0.59%
retprice	0.0076	89.4222	89.4165	-0.01%	87.2661	-2.41%
beer	0.0002	24.2800	24.2735	-0.03%	23.6553	-2.57%
cigsale(1988)	0.0001	90.1000	91.3935	1.44%	113.8237	26.33%
cigsale(1980)	0.9894	120.2000	120.2361	0.03%	138.0895	14.88%
cigsale(1975)	0.0026	127.1000	127.1092	0.01%	136.9316	7.74%

Note: "V.weight" is the optimal covariate weight in the diagonal of V matrix.
"Synthetic Control" is the weighted average of donor units with optimal weights.
"Average Control" is the simple average of all control units with equal weights.

Optimal Unit Weights:

Unit	U.weight
Utah	0.3430
Nevada	0.2410
Montana	0.2180
Colorado	0.1380
Connecticut	0.0600

Note: The unit Alabama Arkansas Delaware Georgia Idaho Illinois Indiana Iowa Kansas Kentucky Louisiana Maine Minnesota Mississippi Missouri Nebraska NewHampshire NewMexico NorthCarolina NorthDakota Ohio Oklahoma Pennsylvania RhodeIsland SouthCarolina SouthDakota Tennessee Texas Vermont Virginia WestVirginia Wisconsin Wyoming in the donor pool get a weight of 0.

Prediction results in the posttreatment periods:

Time	Actual Outcome	Synthetic Outcome	Treatment Effect
1989	82.4000	89.8669	-7.4669
1990	77.8000	87.4213	-9.6213
1991	68.7000	81.9078	-13.2078
1992	67.5000	81.4492	-13.9492
1993	63.4000	81.0495	-17.6495
1994	58.6000	80.6120	-22.0120
1995	56.4000	78.4251	-22.0251
1996	54.5000	77.4402	-22.9402
1997	53.8000	77.7442	-23.9442
1998	52.3000	74.3467	-22.0467
1999	47.2000	73.4932	-26.2932
2000	41.6000	67.2383	-25.6383
Mean	60.3500	79.2495	-18.8995

Note: The average treatment effect over the posttreatment period is -18.8995.

```
Implementing leave-one-out robustness test that excludes one control unit with a
> nonzero weight Utah...Nevada...Montana...Colorado...Connecticut...
```

Leave-one-out robustness test results in the posttreatment period:

Time	Outcome		Synthetic Outcome (LOO)	
	Actual	Synthetic	Min	Max
1989	82.4000	89.8669	88.4724	92.5240
1990	77.8000	87.4213	83.6114	89.0349
1991	68.7000	81.9078	80.9608	82.4889
1992	67.5000	81.4492	80.6239	81.8815
1993	63.4000	81.0495	79.7801	82.1255
1994	58.6000	80.6120	78.6141	83.3754
1995	56.4000	78.4251	75.9901	81.4507
1996	54.5000	77.4402	75.0801	80.6476
1997	53.8000	77.7442	71.7877	84.4861
1998	52.3000	74.3467	71.1668	79.0866
1999	47.2000	73.4932	71.5421	77.5972
2000	41.6000	67.2383	65.0850	70.0086

Note: The last two columns report the minimum and maximum synthetic outcomes when one control unit with a nonzero weight is excluded at a time.

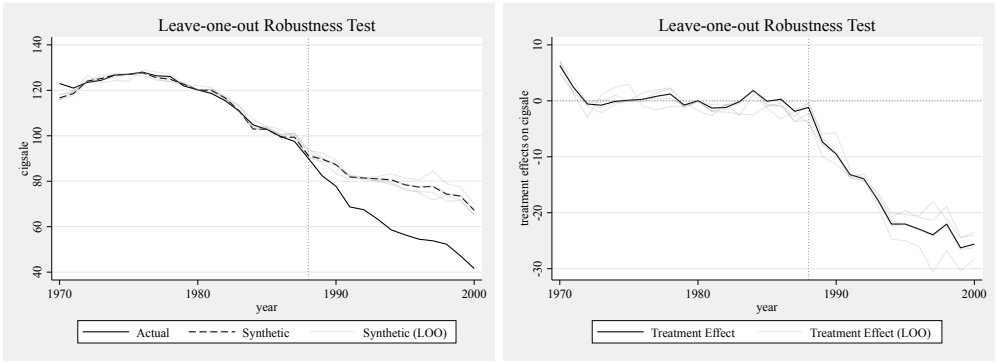
Time	Treatment Effect	Treatment Effect (LOO)	
		Min	Max
1989	-7.4669	-10.1240	-6.0724
1990	-9.6213	-11.2349	-5.8114
1991	-13.2078	-13.7889	-12.2608
1992	-13.9492	-14.3815	-13.1239
1993	-17.6495	-18.7255	-16.3801
1994	-22.0120	-24.7754	-20.0141
1995	-22.0251	-25.0507	-19.5901
1996	-22.9402	-26.1476	-20.5801
1997	-23.9442	-30.6861	-17.9877
1998	-22.0467	-26.7866	-18.8668
1999	-26.2932	-30.3972	-24.3421
2000	-25.6383	-28.4086	-23.4850

Note: The last two columns report the minimum and maximum treatment effects when one control unit with a nonzero weight is excluded at a time.

```
(file california_bias.gph not found)
file california_bias.gph saved
(file california_weight_vars.gph not found)
file california_weight_vars.gph saved
(file california_weight_unit.gph not found)
file california_weight_unit.gph saved
(file california_pred.gph not found)
file california_pred.gph saved
(file california_eff.gph not found)
file california_eff.gph saved
(file california_pred_loo.gph not found)
file california_pred_loo.gph saved
(file california_eff_loo.gph not found)
file california_eff_loo.gph saved
```

Finished.

The above results report the minimum and maximum of predicted outcomes and treatment effects under the LOO scenario, that is, when one of the control units with a nonzero weight is left out in turn. The `synth2` command also produces two graphs for easy inspection, which are collected in figure 4. Figure 4(a) presents the actual outcomes, predicted outcomes, and LOO predicted outcomes. Apparently, the results are qualitatively similar, no matter which control unit with a nonzero weight is excluded. Figure 4(b) graphs the treatment effects and LOO treatment effects. Again, the results appear to be robust in that the estimated treatment effects are not driven by any particular control unit. Note that the LOO robustness test is not a rigorous statistical test, and subjective judgment is sometimes involved in determining the results when the case is not clearly cut.



(a) Actual and predicted paths

(b) Treatment effects

Figure 4. Graphs for the LOO robustness test in example 4

To combine all produced graphs into two columns, we may use the following command:

```
. graph combine `e(graph)', cols(2) altshrink
```

To access the generated Stata frame `california`, we may use the following command:

```
. frame change california
```

To switch back to the default frame containing `smoking.dta`, we can use the following command:

```
. frame change default
```

7 Conclusions

The SCM is a popular method for causal inference in panel data with one treated unit. In this article, we reviewed the SCM methodology and presented the command `synth2`

as a convenient wrapper program for the `synth` command. The `synth2` command provides useful utilities to automate both in-space and in-time placebo tests, as well as the LOO robustness test. Moreover, `synth2` produces a complete set of graphs to visualize the estimation and inference of the SCM. We also demonstrated the use of the `synth2` command by revisiting the classic example of California's tobacco control program (Abadie, Diamond, and Hainmueller 2010). It is our hope that the `synth2` command would free applied researchers from excessive Stata programming and allow them to focus more on substantive research while applying the SCM. Looking forward, as new ways of implementing the SCM and its variants continue to appear, we hope more functionalities may be added to `synth2` or other SCM-related commands.

8 Acknowledgments

The scientific calculation in this article has been done on the HPC Cloud Platform of Shandong University. We thank Hequan Huang for helpful comments. In particular, we are grateful to editor Stephen P. Jenkins and an anonymous referee for very useful comments, and all remaining errors are ours.

9 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 23-3
. net install st0722     (to install program files, if available)
. net get st0722         (to install ancillary files, if available)
```

The command `synth2` can be installed from the Statistical Software Components by typing

```
. ssc install synth2, all replace
```

where the option `all` specifies downloading the example dataset (`smoking.dta`) attached to the `synth2` command and the option `replace` instructs replacement of the previous version of the `synth2` command if installed. Moreover, because the `synth2` command calls on the `synth` command for underlying SCM estimation, one also needs to have the `synth` command installed (if not, use `ssc install synth, replace`). Note that because the `synth` command uses a C++ plugin for numerical optimization, the results might differ slightly on different computers.

10 References

Abadie, A. 2021. Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature* 59: 391–425. https://doi.org/10.1257/jel.20191450.

Abadie, A., A. Diamond, and J. Hainmueller. 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association* 105: 493–505. https://doi.org/10.1198/jasa.2009.ap08746.

———. 2011. synth: Stata module to implement synthetic control methods for comparative case studies. Statistical Software Components S457334, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s457334.html.

———. 2015. Comparative politics and the synthetic control method. *American Journal of Political Science* 59: 495–510. https://doi.org/10.1111/ajps.12116.

Abadie, A., and J. Gardeazabal. 2003. The economic costs of conflict: A case study of the Basque country. *American Economic Review* 93: 113–132. https://doi.org/10.1257/000282803321455188.

Athey, S., and G. W. Imbens. 2017. The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives* 31: 3–32. https://doi.org/10.1257/jep.31.2.3.

Galiani, S., and B. Quistorff. 2017. The synth_runner package: Utilities to automate synthetic control estimation using synth. *Stata Journal* 17: 834–849. https://doi.org/10.1177/1536867X1801700404.

Greathouse, J. 2022. scul: Stata module to implement regularized synthetic control (using LASSO) estimators for single and multiple-treated unit settings. Statistical Software Components S459107, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s459107.html.

Hsiao, C., H. S. Ching, and S. K. Wan. 2012. A panel data approach for program evaluation: Measuring the benefits of political and economic integration of Hong Kong with mainland China. *Journal of Applied Econometrics* 27: 705–740. https://doi.org/10.1002/jae.1230.

Wiltshire, J. C. 2022. allsynth: (Stacked) Synthetic control bias-correction utilities for Stata. Working paper, University of California–Davis. https://justinwiltshire.com/s/allsynth_Wiltshire.pdf.

Yan, G., and Q. Chen. 2022. rcm: A command for the regression control method. *Stata Journal* 22: 842–883. https://doi.org/10.1177/1536867X221140960.

About the authors

Guanpeng Yan is a PhD student at School of Economics, Shandong University, Jinan, Shandong Province, China.

Qiang Chen (corresponding author) is a professor at School of Economics, Shandong University, Jinan, Shandong Province, China.