The Stata Journal (*yyyy*) *vv*, Number *ii*, pp. 1–26

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. Synthetic control method (SCM) (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010) is a popular method for causal inference in panel data with a single treated unit, which often uses placebo tests for statistical inference. While SCM can be implemented by the excellent Stata command `synth`, 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/unit weights, covariate balance, actual/predicted outcomes, treatment effects, placebo tests, post/pre MSPE ratios, pointwise p-values (two-sided, right-sided and left-sided), and the leave-one-out robustness test. We illustrates the use of `synth2` command by revisiting the classic example of California's tobacco control program (Abadie, Diamond, and Hainmueller 2010).

Keywords: st0001, synth2, synth, synthetic control method, placebo test, robustness test, causal inference

1 Introduction

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 a single treated unit. Hailed as "arguably the most important innovation in the policy evaluation literature in the last 15 years" (Athey and Imbens 2017), SCM has spawn a large literature; see Abadie (2021) for an excellent review. Basically, for a treated unit, 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. Due to the small sample sizes often encountered in practice, Abadie, Diamond, and Hainmueller (2010) proposes an in-space placebo test for statistical inference. In addition, Abadie, Diamond, and Hainmueller (2015) recommends an in-time placebo test and a leave-one-out robustness test.

While SCM can be implemented by the excellent Stata command `synth` (Abadie, Diamond, and Hainmueller 2011), it is still uneasy 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 SCM and the case

with many treated units. However, `allsynth` only conducts the in-space placebo test with MSPE-based p-values, 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 regression control method (a.k.a. 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, which are mostly unavailable in Stata until now. First, `synth2` automates the in-space placebo test for SCM (Abadie, Diamond, and Hainmueller 2010), which is previously only available in `synth_runner` and `allsynth` in a limited way. Second, `synth2` implements the in-time placebo test for SCM (Abadie, Diamond, and Hainmueller 2015) for the first time in Stata. Third, `synth2` conducts the leave-one-out robustness test (Abadie, Diamond, and Hainmueller 2015), which is also new in Stata. Last but not least, `synth2` produces a variety of figures for visualization, many of which are previously unavailable in Stata. These include figures to visualize covariate/unit weights, covariate balance, actual/predicted outcomes, treatment effects, placebo tests, post/pre MSPE ratios, pointwise p-values (two-sided, right-sided and left-sided), and the leave-one-out robustness test.

The rest of the paper is organized as follows. Section 2 provides an overview of the methodology of SCM. Section 3 and Section 4 discuss placebo tests and the leave-one-out robustness test respectively. Section 5 presents the Stata 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 Synthetic Control Method

The exposition of 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 $t = 1, \ldots, T_0$ (pre-intervention) and $t = T_0 + 1, \ldots, T$ (post-intervention). 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), which form the "donor pool". Let y_{it}^1 and y_{it}^0 be the outcome of unit i in period t with and without intervention respectively, then the observed outcome y_{it} can be expressed as

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

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, and $\alpha_{it} = y_{it}^1 - y_{it}^0$ denotes the treatment effect for unit i in period t. The goal is to estimate $(\alpha_{1T_0+1}, \ldots, \alpha_{1T})$, which is equivalent to estimating $\left(y_{1T_0+1}^0, \ldots, y_{1T}^0\right)$, since $\left(y_{1T_0+1}^1, \ldots, y_{1T}^1\right)$ 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 (i.e., 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. 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

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

where $\mathbf{w} = (w_2, \ldots, w_{J+1})'$ is 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 $\sum_{i=2}^{N+1} w_i = 1$. SCM selects the optimal $\mathbf{w}$ so that the pretreatment characteristics of the synthetic control are most similar to that of the treated unit. Let $\mathbf{x}_1$ be the $(K \times 1)$ vector containing the pretreatment covariates of the treated unit, and $\mathbf{X}_0$ 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, which 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 $\mathbf{Z}_0$ 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 $\mathbf{V}^*$, which minimizes the mean squared prediction error (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 $\widehat{y}_{1t}^0$ and the treatment effect $\widehat{\Delta}_{1t} = y_{1t} - \widehat{y}_{1t}^0$ over the posttreatment period according to equation (1).

3 Placebo Tests

The conventional statistical inference for SCM relies on placebo tests, which typically come in two forms, i.e., 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 papers by Abadie and coauthors.

3.1 In-space placebo test

The idea of 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 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 since they contain little useful information. The pointwise in-space placebo test considers the following 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", i.e., 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 (i.e., $\widehat{\Delta}_{1t}$ 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 towards 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 towards the left tail of the distribution:

$$H_2 : \Delta_{1t} < 0.$$

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 by 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 ratio of posttreatment MSPE to pretreatment MSPE (denoted as "post/pre MSPE ratio" for short) 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 (i.e., p-value) of obtaining a post/pre MSPE ratio as large as that of the treated unit is calculated as follows,

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

where $\text{MSPE}_{i,post}$ and $\text{MSPE}_{i,pre}$ 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 $\frac{1}{N+1}$.

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 $\left[\widetilde{T}_0, T_0\right]$.

The intuition is that if the estimated placebo effects during the period $\left[\widetilde{T}_0, T_0\right]$ 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 leave-one-out robustness test proposed by Abadie, Diamond, and Hainmueller (2015). As a weighted average of donor units, the optimal synthetic control typically is sparse in the sense that most control units receive

a zero weight. Therefore, one may be concerned that the estimated treatment effects may be disproportionally driven by just a single control unit with a nonzero weight.

The leave-one-out robustness test re-estimates the original synthetic control model by omitting in each iteration one of the original selected donors . Intuitively, the leave-one-out analysis evaluates to what extent results are driven by any particular control unit, although the exclusion of a non-zero-weighted unit sacrifices some goodness of fit. If the outcomes and treatment effects of leave-one-out synthetic controls are similar to those of synthetic control 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(`$\#_c$`)]) loo frame(`*framename*`) savegraph([`*prefix*`],`
`[asis replace]) nofigure]`

`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.

5.2 Required settings

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

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

5.3 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`.

Model

`counit(`*numlist*`)` 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 constitute what is known as the "donor pool". If no `counit` is specified, the donor pool defaults to all available units other than the treated unit.

`preperiod(`*numlist*`)` a list of pretreatment periods as *numlist* given in the time variable specified in `xtset` *timevar*. If no `preperiod` is specified, `preperiod` defaults to the entire pre-intervention period, which ranges from the earliest time period available in the time variable to the period immediately prior to the intervention.

`postperiod(`*numlist*`)` a list of posttreatment periods (when and after the intervention occurred) as *numlist* given in the time variable specified in `xtset` *timevar*. If no `postperiod` is specified, `postperiod` defaults to the entire post-intervention period, which ranges from the time period when the intervention occurred to the latest time period available in the time variable.

`xperiod(`*numlist*`)` 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*`)` a list of pretreatment periods as *numlist* given in the time variable specified in `xtset` *timevar*, over which the mean squared prediction error (MSPE) should be minimized.

`nested` if `nested` is specified, `synth2` embarks 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` if `nested` is specified, the user can also specify `allopt` if she or he is willing to trade-off even more computing time in order to gain fully robust results. `allopt` provides a robustness check by running the nested optimization three times using three different starting points, and returns the best result. For details, see `synth`.

`customV(`*numlist*`)` 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`.

Optimization

`synth2` uses `synth`'s constrained quadratic optimization routine. The options `margin(`*real*`)`,

`maxiter(#)`, `sigf(#)` and `bound(#)` are identical to those of the `synth` command, and the reader is referred to `synth`.

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|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(`*numlist*`)` 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 only applies when `unit` or `unit(`*numlist*`)` is specified. If this option is not specified, then no fake treatment units are discarded.

`show(`$\#_s$`)` specifies the number of units to show in the post/pre MSPE graph, which correponds to units with the largest $\#_s$ ratios of posttreatment MSPE to pretreatment MSPE. This option only applies when `unit` or `unit(`*numlist*`)` is specified. If this option is not specified, the default is to show post/pre MSPE ratios for all units.

Robustness test

`loo` specifies the leave-one-out robustness test that excludes one control unit with a nonzero weight at a time. `synth2` iteratively re-estimates 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 to what extent results are driven by any particular control unit.

Reporting

`frame(`*framename*`)` creates a Stata frame storing 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 already exists, or

created if not.

`nofigure` Do not display figures. The default is to display all figures from estimation results and placebo tests if available.

`savegraph(`[*prefix*]`, `[`asis replace`]`)` automatically and iteratively calls the `graph save` to save all produced graphs to the current path, where *prefix* specifies the prefix added to *_graphname* to form a file name, 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 only applies when `nofigure` is not specified.

5.4 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 fitted in the pretreatment period
`e(r2)`	R^2 of the model fitted 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 $\mathbf{V}$ containing the optimal covariate weights in the diagonal
`e(U_wt)`	vector $\mathbf{w}$ 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

The Stata command `synth2` can be installed from the SSC:

```
. 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 previous version of the `synth2` command if installed. Moreover, since 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 since the `synth` command uses C++ plugin for numerical optimization, the results might differ slightly using different computers.

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). The dataset `smoking.dta` attached to the `synth2` command includes the following variables for 39 US 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 the dataset `smoking.dta`, we declare it as a panel dataset:

```
. use smoking, clear
(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)`", since 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) cigsale(1975)`", which are the values of `cigsale` in 1988, 1980 and 1975 respectively. Moreover, we use the option "`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) ci
> gsale(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:
```

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

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.

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) indicates 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 grey 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.

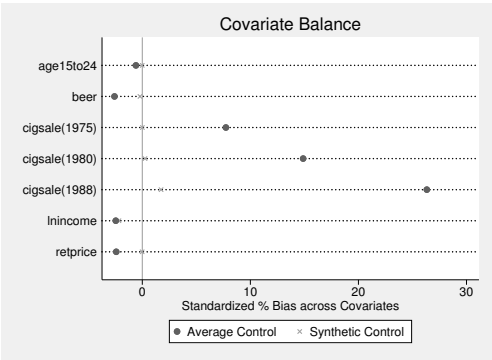
6.2 Example 2: In-space placebo test

In this example, we implement the in-space placebo test. The option "`placebo(unit cut(2))`" is added to request the in-space placebo test using all fake treatment units, but exclude those units with pretreatment MSPE 2 times larger than that of the treated unit. Note that one can also replace "`unit`" with "`unit(`*numlist*`)`" 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 assure convergence. Implementing the following command may be time-consuming, but it is certainly worth the wait.

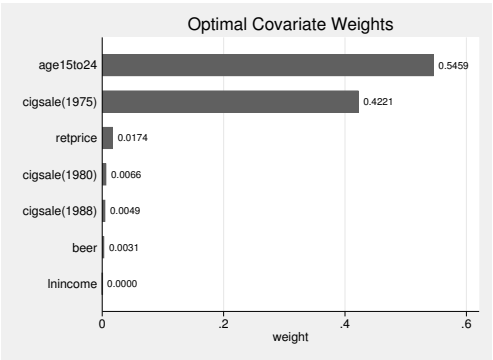
```
. synth2 cigsale lnincome age15to24 retprice beer cigsale(1988) cigsale(1980) ci
> gsale(1975), trunit(3) trperiod(1989) xperiod(1980(1)1988) nested placebo(unit
>  cut(2)) sigf(6)
Fitting results in the pretreatment periods:
```

Treated Unit	: California		Treatment Time	:	1989
Number of Control Units	=	38	Root Mean Squared Error	=	1.77955
Number of Covariates	=	7	R-squared	=	0.97411

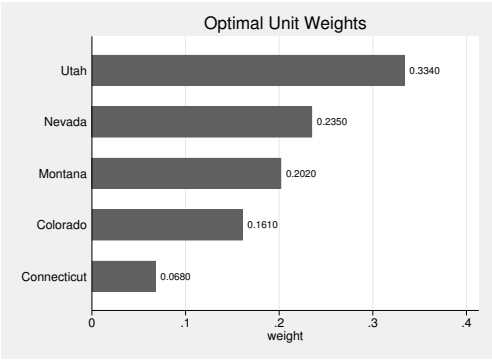
2. To save space, we combine these graphs in a single chart. Commands for retrieving this and other charts containing multiple graphs are provided in the help file and the `example.do` file available from SSC.



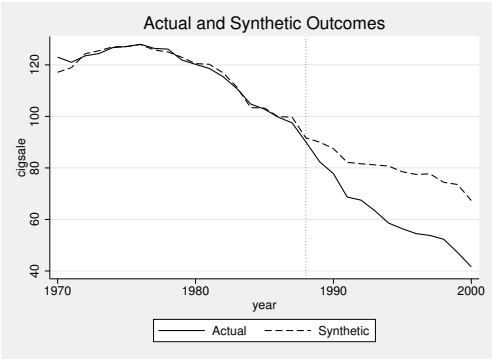
(a) Covariate balance



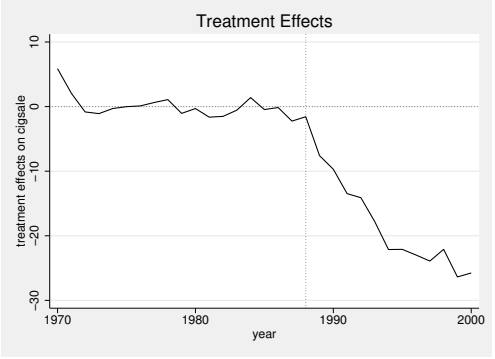
(b) Optimal covariate weights (diagonal of V^*)



(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

Covariate balance in the pretreatment periods:

Covariate	V.weight	Treated	Synthetic Control		Average Control	
			Value	Bias	Value	Bias
lnincome	0.0001	10.0766	9.8527	-2.22%	9.8292	-2.45%
age15to24	0.0020	0.1735	0.1737	0.07%	0.1725	-0.59%
retprice	0.0147	89.4222	89.3766	-0.05%	87.2661	-2.41%
beer	0.0083	24.2800	24.2236	-0.23%	23.6553	-2.57%
cigsale(1988)	0.0142	90.1000	91.3866	1.43%	113.8237	26.33%
cigsale(1980)	0.1923	120.2000	120.2357	0.03%	138.0895	14.88%
cigsale(1975)	0.7683	127.1000	127.0999	-0.00%	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	0.3450
Nevada	0.2410
Montana	0.2070
Colorado	0.1480
Connecticut	0.0590

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	89.8201	-7.4201
1990	77.8000	87.3789	-9.5789
1991	68.7000	81.9182	-13.2182
1992	67.5000	81.4061	-13.9061
1993	63.4000	81.0228	-17.6228
1994	58.6000	80.5678	-21.9678
1995	56.4000	78.3083	-21.9083
1996	54.5000	77.3429	-22.8429
1997	53.8000	77.6174	-23.8174
1998	52.3000	74.1877	-21.8877
1999	47.2000	73.3950	-26.1950
2000	41.6000	67.1478	-25.5478
Mean	60.3500	79.1761	-18.8261

Note: The average treatment effect over the posttreatment period is -18.8261.

```
Implementing placebo test using fake treatment unit Alabama...Arkansas...Colorad
> o...Connecticut...Delaware...Georgia...Idaho...Illinois...Indiana...Iowa...Kan
> sas...Kentucky...Louisiana...Maine...Minnesota...Mississippi...Missouri...Mont
> ana...Nebraska...Nevada...NewHampshire...NewMexico...NorthCarolina...NorthDako
```

```
> ta...Ohio...Oklahoma...Pennsylvania...RhodeIsland...SouthCarolina...SouthDakot
> a...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	3.1668	391.2533	123.5490	1.0000
Alabama	5.4170	8.2108	1.5157	1.7106
Arkansas	4.5587	28.8239	6.3228	1.4395
Colorado	17.6103	68.6736	3.8996	5.5609
Connecticut	20.6396	118.9189	5.7617	6.5175
Delaware	30.3949	499.0694	16.4195	9.5980
Georgia	1.4610	116.8893	80.0074	0.4613
Idaho	5.8142	39.1830	6.7392	1.8360
Illinois	4.3146	89.0552	20.6406	1.3624
Indiana	14.4145	469.4150	32.5654	4.5518
Iowa	14.6527	31.5816	2.1553	4.6270
Kansas	14.1121	9.0349	0.6402	4.4563
Kentucky	431.7229	1475.7975	3.4184	136.3284
Louisiana	2.0183	94.1070	46.6279	0.6373
Maine	7.7461	52.9939	6.8413	2.4461
Minnesota	14.1736	52.5217	3.7056	4.4757
Mississippi	4.0894	37.2754	9.1151	1.2913
Missouri	1.2009	85.1794	70.9308	0.3792
Montana	5.2861	54.8978	10.3853	1.6692
Nebraska	4.8287	36.5597	7.5713	1.5248
Nevada	40.6500	83.4186	2.0521	12.8364
NewHampshire	3436.5980	134.9018	0.0393	1085.2007
NewMexico	5.0577	63.7459	12.6036	1.5971
NorthCarolina	90.2241	67.3144	0.7461	28.4907
NorthDakota	8.0725	72.3200	8.9588	2.5491
Ohio	2.9585	12.3535	4.1757	0.9342
Oklahoma	5.7128	267.8078	46.8786	1.8040
Pennsylvania	2.6308	6.1299	2.3301	0.8307
RhodeIsland	87.9904	242.6697	2.7579	27.7854
SouthCarolina	2.1997	41.2941	18.7727	0.6946
SouthDakota	7.5688	32.2367	4.2592	2.3901
Tennessee	5.2043	123.3097	23.6938	1.6434
Texas	4.6691	239.8559	51.3707	1.4744
Utah	593.7643	223.2758	0.3760	187.4975
Vermont	15.3860	116.8473	7.5944	4.8585
Virginia	2.7825	219.8136	78.9994	0.8786
WestVirginia	8.1492	242.1734	29.7175	2.5733
Wisconsin	2.8950	75.2425	25.9901	0.9142
Wyoming	83.7717	31.8266	0.3799	26.4532

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 0.0500.
(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 19 units with pretreatment MSPE 2 times larger than the treated unit, including Colorado Connecticut Delaware 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	-7.4201	0.0500	1.0000	0.0500
1990	-9.5789	0.1000	0.9500	0.1000
1991	-13.2182	0.1500	0.9000	0.1500
1992	-13.9061	0.1000	0.9500	0.1000
1993	-17.6228	0.0500	1.0000	0.0500
1994	-21.9678	0.0500	1.0000	0.0500
1995	-21.9083	0.0500	1.0000	0.0500
1996	-22.8429	0.0500	1.0000	0.0500
1997	-23.8174	0.0500	1.0000	0.0500
1998	-21.8877	0.1000	0.9500	0.1000
1999	-26.1950	0.0500	1.0000	0.0500
2000	-25.5478	0.0500	1.0000	0.0500

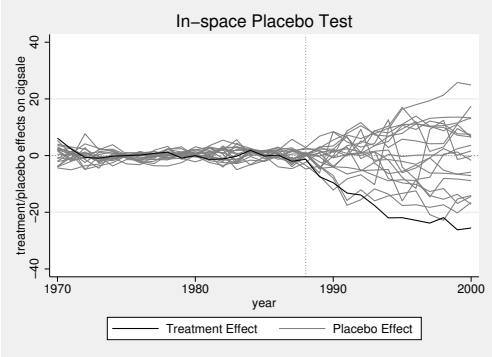
```
Note: (1) The two-sided p-value of the treatment effect for a particular
      period is defined as the frequency that the absolute values of the
      placebo effects are greater than or equal to the absolute value of
      treatment effect.
      (2) The right-sided (left-sided) p-value of the 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 MSPE two times larger than the treated unit, the MSPE-based p-value is still 0.05. 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.

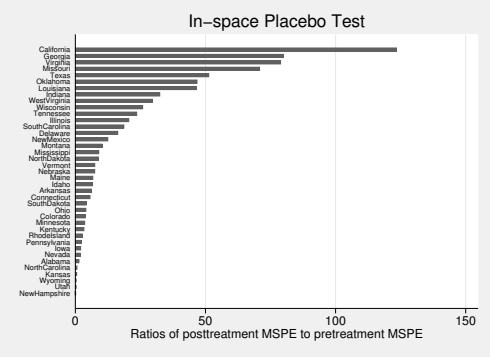
In the meantime, the above `synth2` command produces five graphs collected in Figure 2. Figure 2(a) graphs the distribution of placebo effects, against which the estimated treatment effects are compared. Apparently, the estimated treatment effects are all negative, and mostly lie at the bottom of the 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(`$\#_s$`)` to restrict the number of units to display in this graph; such as `placebo(unit cut(2) show(10))`. Figure 2(c), 2(d) and 2(e) graph two-sided, right-sided and left-sided p-values respectively.

6.3 Example 3: In-time placebo test

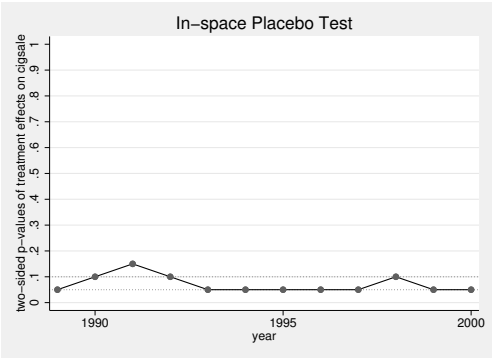
In this example, we implement the in-time placebo test. The option "`placebo(period (1985))`" specifies the in-time placebo test with 1985 as the fake treatment time, which is 4 years earlier than the actual treatment time of 1989. In addition, we remove the



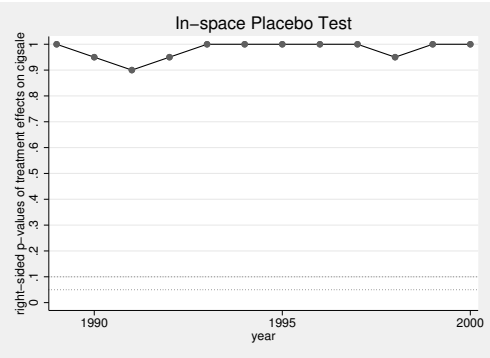
(a) Treatment and placebo effects



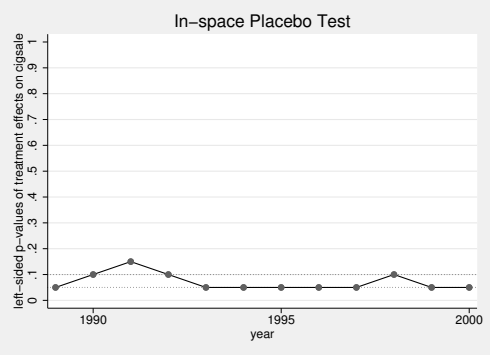
(b) Post/pre MSPE ratios



(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

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)` by `cigsale(1984)`, which are unreported to save space.

```
. synth2 cigsale lnincome age15to24 retprice beer cigsale(1980) cigsale(1975), t
> runit(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	= 2.20530
Number of Covariates	= 6	R-squared	= 0.95253

Covariate balance in the pretreatment periods:

Covariate	V.weight	Treated	Synthetic Control		Average Control	
			Value	Bias	Value	Bias
lnincome	0.0000	10.0372	9.8639	-1.73%	9.7892	-2.47%
age15to24	0.0000	0.1815	0.1825	0.55%	0.1814	-0.06%
retprice	0.0384	76.2200	76.1523	-0.09%	71.8353	-5.75%
beer	0.0000	25.0000	23.0089	-7.96%	23.6947	-5.22%
cigsale(1980)	0.9597	120.2000	120.0846	-0.10%	138.0895	14.88%
cigsale(1975)	0.0019	127.1000	126.8324	-0.21%	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	0.0500

```
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	93.0322	-10.6322
1990	77.8000	89.4297	-11.6297
1991	68.7000	82.4727	-13.7727
1992	67.5000	80.6731	-13.1731
1993	63.4000	79.5929	-16.1929
1994	58.6000	78.1272	-19.5272
1995	56.4000	75.6207	-19.2207

1996	54.5000	74.8372	-20.3372
1997	53.8000	74.5395	-20.7395
1998	52.3000	71.1561	-18.8561
1999	47.2000	71.4380	-24.2380
2000	41.6000	65.8382	-24.2382
Mean	60.3500	78.0631	-17.7131

```
Note: The average treatment effect over the posttreatment period is -17.7131.
Implementing placebo test using fake treatment time 1985...
In-time placebo test results using fake treatment time 1985:
```

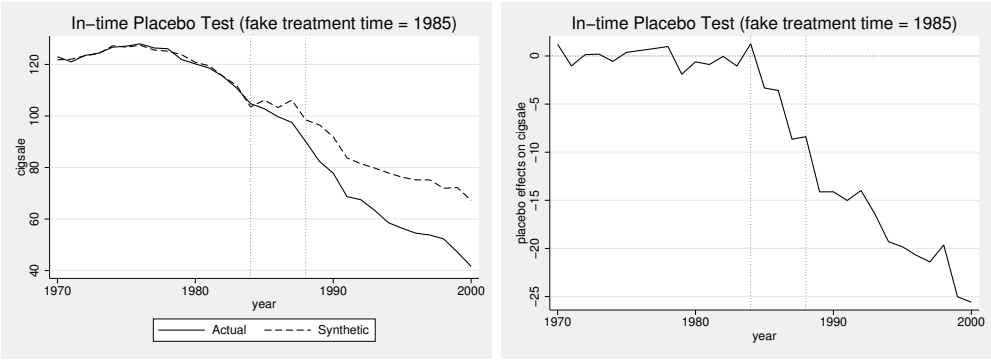
Time	Actual Outcome	Synthetic Outcome	Treatment Effect
1985	102.8000	106.1262	-3.3262
1986	99.7000	103.2850	-3.5850
1987	97.5000	106.1524	-8.6524
1988	90.1000	98.4873	-8.3873
1989	82.4000	96.5237	-14.1237
1990	77.8000	91.9127	-14.1127
1991	68.7000	83.7156	-15.0156
1992	67.5000	81.4730	-13.9730
1993	63.4000	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 the treatment starting from 1985. There appears 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 anti-smoking movement might have started a few years earlier in California, which culminated in the passage of Proposition 99 in 1988.

6.4 Example 4: Leave-one-out robustness test

In this example, we implement the leave-one-out 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 (e.g., to



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

Figure 3: Graphs for the in-time placebo test in Example 3

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) ci
> gsale(1975), trunit(3) trperiod(1989) xperiod(1980(1)1988) nested loo frame(ca
> lifornia) savegraph(california, replace)
Fitting results in the pretreatment periods:
```

Treated Unit	: California	Treatment Time	: 1989
Number of Control Units	= 38	Root Mean Squared Error	= 1.78329
Number of Covariates	= 7	R-squared	= 0.97365

Covariate balance in the pretreatment periods:

Covariate	V.weight	Treated	Synthetic Control		Average Control	
			Value	Bias	Value	Bias
lnincome	0.0002	10.0766	9.8509	-2.24%	9.8292	-2.45%
age15to24	0.0124	0.1735	0.1736	0.03%	0.1725	-0.59%
retprice	0.0113	89.4222	89.3331	-0.10%	87.2661	-2.41%
beer	0.0416	24.2800	24.2554	-0.10%	23.6553	-2.57%
cigsale(1988)	0.0231	90.1000	91.2783	1.31%	113.8237	26.33%
cigsale(1980)	0.4378	120.2000	120.1872	-0.01%	138.0895	14.88%
cigsale(1975)	0.4735	127.1000	126.9907	-0.09%	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	0.3420
Nevada	0.2380
Montana	0.2170
Colorado	0.1450
Connecticut	0.0580

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	89.7304	-7.3304
1990	77.8000	87.3001	-9.5001
1991	68.7000	81.8829	-13.1829
1992	67.5000	81.4287	-13.9287
1993	63.4000	81.0450	-17.6450
1994	58.6000	80.6229	-22.0229
1995	56.4000	78.4191	-22.0191
1996	54.5000	77.4316	-22.9316
1997	53.8000	77.7288	-23.9288
1998	52.3000	74.3255	-22.0255
1999	47.2000	73.4654	-26.2654
2000	41.6000	67.2107	-25.6107
Mean	60.3500	79.2159	-18.8659

Note: The average treatment effect over the posttreatment period is -18.8659.

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	89.7304	88.3892	92.3509
1990	77.8000	87.3001	83.5373	89.2205
1991	68.7000	81.8829	80.8905	82.4889
1992	67.5000	81.4287	80.6239	81.8815
1993	63.4000	81.0450	79.7801	82.0592
1994	58.6000	80.6229	78.6141	83.3112
1995	56.4000	78.4191	75.9901	81.3864
1996	54.5000	77.4316	75.0801	80.5833
1997	53.8000	77.7288	71.7877	84.4150
1998	52.3000	74.3255	71.1668	79.0314
1999	47.2000	73.4654	71.5421	77.5396
2000	41.6000	67.2107	65.0850	69.9503

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	-7.3304	-9.9509	-5.9892

1990	-9.5001	-11.4205	-5.7373
1991	-13.1829	-13.7889	-12.1905
1992	-13.9287	-14.3815	-13.1239
1993	-17.6450	-18.6592	-16.3801
1994	-22.0229	-24.7112	-20.0141
1995	-22.0191	-24.9864	-19.5901
1996	-22.9316	-26.0833	-20.5801
1997	-23.9288	-30.6150	-17.9877
1998	-22.0255	-26.7314	-18.8668
1999	-26.2654	-30.3396	-24.3421
2000	-25.6107	-28.3503	-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 minima and maxima of predicted outcomes and treatment effects under the leave-one-out scenario, i.e., 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, as well as leave-one-out (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 leave-one-out (LOO) treatment effects. Again, the results appear to be robust in the sense that the estimated treatment effects are not driven by any particular control unit. Note that the leave-one-out 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.

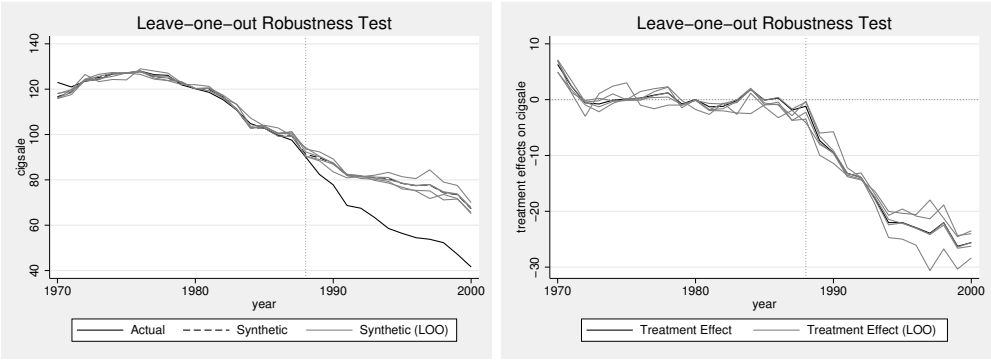
To combine all produced graphs into two columns, we may use the following command:

```
. 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 the dataset `smoking.dta`, we can use the following command:



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

Figure 4: Graphs for the leave-one-out robustness test in Example 4

```
. frame change default
```

7 Conclusions

Synthetic control method (SCM) is a popular method for causal inference in panel data with a single treated unit. In this paper, we review the SCM methodology, and present the Stata 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 leave-one-out robustness test. Moreover, `synth2` produces a complete set of graphs to visualize the estimation and inference of SCM. We also demonstrate the use of `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 SCM. Looking forward, as new ways of implementing SCM and its variants continue to appear, more functionalities may be added to `synth2` or other SCM-related commands.

8 Acknowledgments

The scientific calculation in this paper 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 References

Abadie, A. 2021. Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature* 59(2): 391–425.

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(490): 493–505.

———. 2011. SYNTH: Stata module to implement Synthetic Control Methods for Comparative Case Studies. Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s457334.html.

———. 2015. Comparative politics and the synthetic control method. *American Journal of Political Science* 59(2): 495–510.

Abadie, A., and J. Gardeazabal. 2003. The economic costs of conflict: A case study of the Basque Country. *American Economic Review* 93(1): 113–132.

Athey, S., and G. W. Imbens. 2017. The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives* 31(2): 3–32.

Galiani, S., and B. Quistorff. 2017. The synth_runner package: Utilities to automate synthetic control estimation using synth. *The Stata Journal* 17(4): 834–849.

Greathouse, J. 2022. SCUL: Stata module to implement regularized synthetic control (using LASSO) estimators for single and multiple-treated unit settings. Statistical Software Components, Boston College Department of Economics. 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(5): 705–740.

Wiltshire, J. C. 2022. allsynth: (Stacked) Synthetic Control Bias-Correction Utilities for Stata. *Working Paper* . https://justinwiltshire.com/s/allsynth_Wiltshire.pdf.

Yan, G., and Q. Chen. 2022. rcm: A Stata Command for Regression Control Method. *Stata Journal* .

About the authors

Guanpeng Yan is a Ph.D. 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.