Causal Analysis

Impact Evaluation and Causal Machine Learning with Applications in R

Chapter 8: Synthetic Controls

Table of Contents

Causal Analysis with One Treated Unit

Synthetic control method (Abadie and Gardeazabal, 2003; Abadie, Diamond, and Hainmueller, 2011):

- Requires that the outcome is observed in both pretreatment and posttreatment periods.
- Requires panel data, such that the same subjects can be followed over time.
- Originally developed for case study settings to evaluate the treatment effect on a single treated unit when constructing a nontreated comparison observation based on a pool of multiple nontreated units.

Synthetic Control Method - Example

Example

- Synthetic control study by Abadie and Gardeazabal (2003) on the impact of terrorist conflict on Basque Country's GDP per capita.
- Treated unit is the Basque Country while synthetic control is created from other Spanish regions.

Graphical illustration (Figure 1):

- Treatment starts in 1968, marked by ETA's first victim.
- Solid line: Basque Country's GDP per capita.
- Dashed line: Synthetic control's GDP per capita.
- Post-1968: Divergence in GDP per capita, indicating a negative treatment effect.

Graphical Illustration

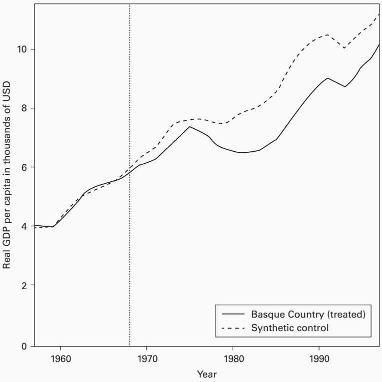

Figure 1: Synthetic control method applied to evaluate terrorist activities in the Basque Country.

Notation

- Panel data of n units, indexed by $i \in \{1, \dots, n\}$.
- Observed over $\mathcal{T}$ time periods, indexed by $t \in \{1, \dots, \mathcal{T}\}$.
- Y_{it}: Observed outcome of unit i in period t.
- Only the last unit $i = n$ is treated in period $T_0 + 1$, where T_0 denotes the last period prior to treatment such that $T_0 \geq 1$.

Treatment Effect Estimation

- Treatment effect estimate for treated unit $i = n$ for any posttreatment period $t \geq T_0 + 1$ is the difference between the treated outcome and a weighted average of nontreated outcomes in that period. Formally:

$$\hat{\Delta}_{n,T=t} = Y_{nt} - \sum_{i=1}^{n-1} \hat{\omega}_i Y_{it}, \text{ for any } t \geq T_0 + 1, \tag{8.1}$$

  where $\hat{\omega}_i$ is a specific weight (importance) of a nontreated unit.

- Weights are chosen such that the weighted average of pretreatment outcomes of nontreated units matches the development of the pretreatment outcome of the treated unit, up to period T_0. Formally, choose weights such that:

$$\sum_{i=1}^{n-1} \hat{\omega}_i Y_{it} \approx Y_{nt}, \text{ for all } t = 1, \ldots, T_0. \tag{8.2}$$

Identification Assumptions

- Controlling for pretreatment outcomes is sufficient to control for confounders that entail diverging potential outcomes under nontreatment of the treated and nontreated units in the posttreatment periods (related to selection on observables).
- Further assumptions: no anticipation assumption and the convex hull condition (a common support assumption).
- Convex hull condition requires that the pretreatment outcomes of the treated unit are not too extreme compared to the nontreated units (not much higher or lower than the highest or lowest outcome of nontreated units in any pretreatment period).

Calculation of Weights

- Weights $\hat{\omega}_i$ are either positive or zero: $\hat{\omega}_i \geq 0$.
- Weights sum up to 1: $\sum_{i=1}^{n-1} \hat{\omega}_i = 1$.
- Computed using least squares approach:

$$\hat{\omega} = \arg\min_{\omega^*} \sum_{t=1}^{T_0} (Y_{nt} - \omega_1^* Y_{1t} - \ldots - \omega_{n-1}^* Y_{n-1t})^2,$$
$$\text{subject to } \omega_i^* \geq 0, \sum_{i=1}^{n-1} \omega_i^* = 1. \tag{8.3}$$

- Allowing for period-dependent weights $\nu_t \geq 0$ in addition:

$$\hat{\omega} = \arg\min_{\omega^*} \sum_{t=1}^{T_0} \nu_t (Y_{nt} - \omega_1^* Y_{1t} - \ldots - \omega_{n-1}^* Y_{n-1t})^2,$$
$$\text{subject to } \omega_i^* \geq 0, \sum_{i=1}^{n-1} \omega_i^* = 1. \tag{8.4}$$

Transformed Outcomes

- One may transform outcomes prior to running the synthetic control method.
- Taking logarithms yields a causal effect which can be interpreted in terms of percentage change.
- Demeaned outcomes $\tilde{Y}_{it}$:

$$\tilde{Y}_{it} = Y_{it} - \frac{1}{T_0} \sum_{t=1}^{T_0} Y_{it} \tag{8.5}$$

- Considering $\tilde{Y}_{it}$ instead of Y_{it} implies that the weighting aims for a combination of nontreated units with pretreatment outcome trends similar to the treated unit, instead of similar levels.
- Hence, we may implement the synthetic control method also based on common-trend-type assumptions.

Synthetic Control Method with Covariates

Including covariates may make the selection-on-observables-type (or common trend in the case of demeaned outcomes) assumption more plausible.

Multiple ways to include them:

- Include covariates in the optimization problem (8.3) similar to pretreatment outcomes.
- Regress outcomes on covariates, and then use outcome residuals in the optimization problem (8.3), as suggested by Doudchenko and Imbens (2016).

Inference for the Synthetic Control Method (1)

Determining statistical significance of treatment effects is challenging due to having only a single treated unit.

Randomized inference based on permutation, see Abadie, Diamond, and Hainmueller, (2010):

- Iteratively consider one nontreated unit as treated and the other nontreated units as nontreated to estimate a placebo effect.
- Distribution of placebo effects allows evaluating how extreme the actual treatment effect is relative to placebo effects.
- Approximate p-value as the share of placebo effects which are more extreme than the treatment effect.

Inference for the Synthetic Control Method (2)

Further inference approaches:

- Conformal inference, see Chernozhukov, Wüthrich, and Zhu (2021).
- Based on reassigning pretreatment placebo effects on the treated unit to posttreatment periods and vice versa, in order to compute test statistics based on these permutations.
- Asymptotic variance approximation, see Li (2020).
- Suitable for inference on mean effects across all posttreatment periods, if they are sufficiently numerous.

Table of Contents

Synthetic Control Method and Ordinary Least Squares

- Synthetic control method is related to ordinary least squares (OLS) regression.
- OLS solves the following minimization problem:

$$(\hat{\omega}, \hat{\alpha}) = \arg \min_{\omega^*, \alpha^*} \sum_{t=1}^{T_0} (Y_{nt} - \alpha^* - \omega_1^* Y_{1t} - \ldots - \omega_{n-1}^* Y_{n-1t})^2. \tag{8.6}$$

- Imposing $\omega_i^* \geq 0, \sum_{i=1}^{n-1} \omega_i^* = 1,$ and $\alpha^* = 0$ yields the synthetic control approach in equation (8.3).
- Allowing for a non-zero constant α^* while keeping the conditions on the weights, yields the DiD-related synthetic control approach.
- Dropping the conditions on the weights permits the weights in equation (8.6) to be negative, and thus the method might extrapolate and make predictions beyond the convex hull.

Machine Learning Extensions

Lasso regression:

- Add a penalty term $\lambda \sum_{i=1}^{n-1} |\omega_i^*|$ to equation (8.6).
- Apply cross-validation by repeatedly estimating placebo treatment effects with different penalty terms.
- Select the penalty term that minimizes the mean squared placebo treatment effects (Doudchenko and Imbens, 2016).

Constrained lasso:

- Add constraint on weights $\sum_{i=1}^{n-1} |\omega_i^*| \leq 1$ (Raskutti, Wainwright, and Yu, 2011), which does not rely on cross-validation.

Combination of multiple approaches (Abadie and L'Hour, 2018):

- Apply cross-validation to determine the weight given to each method, which perform better than using a single approach.

Synthetic Control with Multiple Treated Units (1)

- Synthetic control method can be extended to allow for multiple treated units instead of a single treated unit.
- Permits estimating the ATET in a specific outcome period: $\Delta_{D=1,T=t} = E[Y_t(1) - Y_t(0)|D = 1]$.
- One approach is to apply the estimator separately to each of the treated units, and then average over the effects to estimate the ATET:

$$\hat{\Delta}_{D=1,T=t} = \frac{1}{\sum_{i=1}^{n} D_i} \sum_{i:D_i=1} \hat{\Delta}_{i,T=t}. \tag{8.7}$$

  with $\hat{\Delta}_{i,T=t}$ being the effect estimate of the treated unit i in outcome period t.
- For each of the treated unit i, weights are chosen such that equation (8.2) holds.
- Estimate of each treated unit, $\hat{\Delta}_{i,T=t}$, is based on equation (8.1).

Synthetic Control with Multiple Treated Units (2)

- Alternative: compute the weights such that equation (8.2) holds on average for treated units (rather than for each treated unit).
- Denote by $n_1 = \sum_{i=1}^{n} D_i$the number of treated units and assume that the $n - n_1$ nontreated units appear at the top of the data.
- Modify equation (8.2) as follows:

$$\sum_{i=1}^{n-n_1} \hat{\omega}_i Y_{it} + \hat{\alpha} \approx \frac{1}{n_1} \sum_{j=n-n_1+1}^{n} Y_{jt}, \text{ for all } t = 0, \ldots, T_0, \tag{8.8}$$

DiD-type synthetic control approach if $\hat{\alpha}$ can be non-zero.

- Equation (8.3) becomes:

$$(\hat{\omega}, \hat{\alpha}) = \arg \min_{\omega^*, \alpha^*} \sum_{t=1}^{T_0} \left(\frac{1}{n_1} \sum_{j=n-n_1+1}^{n} Y_{jt} - \alpha^* - \omega_1^* Y_{1t} - \ldots - \omega_{n-n_1}^* Y_{n-n_1 t} \right)^2,$$

$$\text{subject to } \omega_i^* \geq 0, \sum_{i=1}^{n-1} \omega_i^* = 1. \tag{8.9}$$

Synthetic Control with Multiple Treated Units (3)

- Arkhangelsky et al. (2019) provide a synthetic DiD approach related to equation (8.9), which in addition incorporates reweighting periods (see equation (8.4)) in a data-driven way.
- As synthetic control/DiD methods can also be applied to multiple treatments, they provide an alternative estimation strategy to the methods in chapters 4 and 7 under the selection-on-observables and common trend assumptions.
- Also under multiple treated units, synthetic control/DiD methods can be combined with other (e.g., regression) methods, with the weights of each method determined by cross-validation.
- Methods can be adapted to staggered treatment introduction, see Ben-Michael, Feller, and Rothstein (2021b).