

Impact Evaluation in Firms and Organizations

With Applications in R and Python

Chapter 9: Synthetic Controls

9.1 Impact Evaluation When a Single Unit Receives the Intervention

9.2 Behavioral Assumptions and Variants

Settings with only one treated unit

- Interventions may affect only a single individual or region rather than several or many
- Synthetic control is designed for single-unit treatment scenarios

Requirements

- Requires panel data, i.e., same units observed over time
- Requires pre- and postintervention outcomes for the same units
- Differs from difference-in-differences, which also works with repeated cross-sectional data, i.e., different units over time
- Synthetic control is suited for situations when only one unit receives the intervention while difference-in-differences is not

Example

New pricing policy

- New pricing policy introduced in one specific market
- Annual sales observed for all markets over multiple years
- Directly observing a truly comparable untreated market may be difficult or even impossible

Role of the synthetic control approach

- Constructs a comparison unit that is a weighted combination of multiple untreated markets
- Comparison unit matches the preintervention sales evolution of the treated market, providing a counterfactual trajectory without intervention

Constructing the Synthetic Control

Synthetic counterfactual

- Counterfactual outcome for the treated market is constructed synthetically
- Create a weighted average of outcomes from control markets such that markets more similar in preintervention sales receive higher weights
- Ensures credible comparison of postintervention outcomes

Impact evaluation

- Assume synthetic control reflects untreated trajectory of the targeted market
- Difference between postintervention treated and synthetic control outcomes provides estimated impact of the new pricing policy

Synthetic Control Method

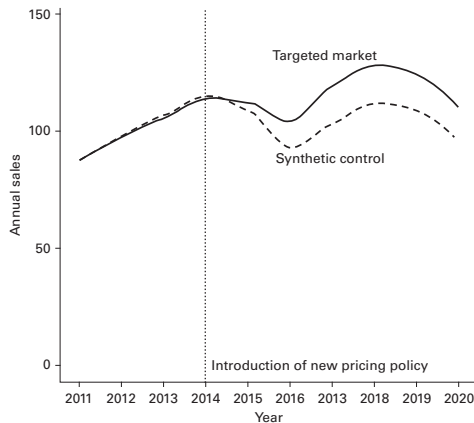


Figure 7.1: Synthetic control method

9.1 Impact Evaluation When a Single Unit Receives the Intervention

9.2 Behavioral Assumptions and Variants

Required Data Structure

Units and time periods

- Dataset contains n units indexed by $i \in \{1, \dots, n\}$
- Each unit is observed over $\mathcal{T}$ periods indexed by $t \in \{1, \dots, \mathcal{T}\}$
- Y_{it} denotes the outcome of unit i in period t

Treatment assignment

- Only one unit $i = n$ is exposed to the intervention
- T_0 represents the last preintervention period and satisfies $T_0 \geq 1$
- Treatment begins in period $T_0 + 1$

Panel data setting

- Outcomes are observed for all units across all periods
- Panel structure allows comparison of pre- and postintervention trajectories

Impact estimation

- Intervention impact is assessed for the treated unit $i = n$ in postintervention periods
- Compare observed outcome of the treated unit to the weighted outcomes of the control units
- Weights are applied to outcomes of units 1 to $n - 1$
- Intervention impact in period t is formally expressed as

$$Y_{nt} - (\hat{\omega}_1 Y_{1t} + \hat{\omega}_2 Y_{2t} + \hat{\omega}_3 Y_{3t} + \cdots + \hat{\omega}_{n-1} Y_{n-1,t}) \quad (9.1)$$

- This equation applies for all postintervention periods $t \geq T_0 + 1$

Synthetic Control Weights

Definition of the synthetic control

- Synthetic control represents the counterfactual outcome for the treated unit
- Computed as a weighted average of control group outcomes

Role of the weights

- Weights $\hat{w}_1, \dots, \hat{w}_{n-1}$ multiply each control unit's outcome
- Weights capture similarity to the treated unit in preintervention periods

Weight properties

- Weights are nonnegative and sum to 100%
- Only selected control units may receive positive weights
- Others may receive zero weight if not comparable enough

Rationale for matching

- Preintervention outcomes are used as characteristics to make the targeted unit and the control group comparable
- Ensures valid comparison of actual and synthetic postintervention trajectories

Weight selection criterion

- Weights chosen so the synthetic control approximates the treated unit before treatment
- Choose weights such that

$$(\hat{\omega}_1 Y_{1t} + \hat{\omega}_2 Y_{2t} + \hat{\omega}_3 Y_{3t} + \cdots + \hat{\omega}_{n-1} Y_{n-1,t}) \approx Y_{nt} \quad \text{for all } t \leq T_0 \quad (9.2)$$

- Approximate equality ensures close preintervention match

Convex hull condition

- Requires preintervention outcomes of the treated unit to lie within the range of control units
- For example, sales in treated unit must not be much higher or lower than in any control unit
- Ensures existence of comparable control units with similar pretreatment outcomes

No anticipation requirement

- Preintervention potential outcomes of the targeted unit must not be affected by future intervention
- For example, sales should not react in advance to the upcoming pricing policy
- Ensures pretreatment periods reflect untreated behavior

Extensions of the Synthetic Control Method

Including additional characteristics

- Characteristics X may be included to make synthetic control resemble more the targeted unit
- Ensures similarity in both preintervention outcomes and observed covariates

Difference-in-differences-based synthetic control

- Synthetic control can instead match trends rather than levels in preintervention periods
- Focuses on similarity in outcome changes instead of levels

Choosing between approaches

- Synthetic difference-in-differences selects best-performing method in a data-driven way using preintervention fit
- Allows flexible choice between difference-in-differences-based and selection-on-observables-based synthetic controls

Effects Based on the Synthetic DiD Method

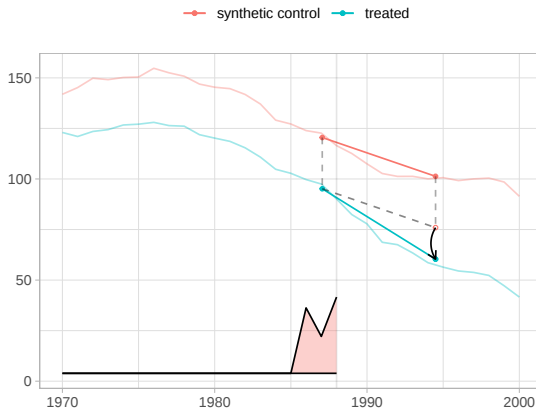


Figure 7.2: Effects based on the synthetic DiD method

Permutation Testing for Inference

Placebo-based inference

- Conventional inference not applicable with single treated unit
- Use permutation testing to approximate significance by estimating placebo effects

Placebo effects

- Each control unit assigned a pseudo-intervention with true impact zero
- Remaining control units form the synthetic control group
- Produces distribution of placebo effects for comparison

Interpreting results

- Compare actual impact to distribution of placebo effects
- Large gap between actual and placebo effects suggests nonzero impact

Limits of permutation testing

- Validity requires random assignment of the intervention
- Synthetic control settings typically lack random assignment

Conformal inference approach

- Focuses on placebo effects of the treated unit rather than control units
- Compares preintervention placebo effects with postintervention effects

Permutation under conformal inference

- Reassign preintervention placebo effects to postintervention periods and vice versa to generate permutation-based effect distributions
- Large gap between actual and permutation-based effects suggests nonzero impact