Synthetic Control Method
Econ 672

The Takeaway

- A generalized difference-in-differences method that can get around the problem of having only 1 treatment group cluster.
- This method uses predictive covariates to generate a synthetic control group for the treatment group by applying an optimal set of weights
- We are mixing predictive analytics and causal inference at an aggregate level

Pros

- Synthetic Control can deal with having only 1 treatment group cluster where as in canonical 2-by-2 DD has inconsistent standard error
- It is visually intensive like RDD
- Uses weighted average of all comparison units to generate a synthetic control group based on characteristics of the treatment unit(s)
    - Weights are transparent in how they are generated
- Helps prevent ad hoc and subjective selection of the control group – prevents peeking at the results
- Helps prevent standard errors from reflecting the sample variance instead of the ability of the control group to reproduce the counterfactual
- Bridges the gap qualitative and quantitative studies

Cons

- Generating weights requires matching on observable predictive characteristics and we cannot use unobserved predictive covariates
- Pre-treatment trends are needed similar to parallel trends in 2-by-2 DD
- Inference of exact p-values needs to be calculated

Assumptions

- Pre-treatment trends
    - Pre-treatment trends need to be satisfied similar to the parallel trend assumptions in 2-by-2 DD
- Convex Hull Assumption
    - This assumption states that Y_{1it} is in the convex hull or minimum set or space that contain $Y_{it}, i \neq j$
    - In other workers, the treatment unit has outcomes similar to other comparison units before treatment (very similar to pre-treatment trends)
    - If this assumption fails, then the treatment unit is not comparable to comparison units

Testable Assumptions

- We can directly test the pre-treatment trends is similar between the treatment unit and comparison units
- We indirectly test the convex hull assumption that outcomes are similar before treatment

The Estimator

- The causal parameter of interest for the synthetic control method is a comparison between the treatment group in time t compared to a weighted average of the synthetic outcome in time t

- $\hat{\delta}_{it} = Y_{1t} - \sum_{2}^{J+1} w_j^*(V) Y_{jt}$
- Where
  - Our optimal weights $w_j^*(V)$ are a function of V, which are the imporantace weights that depend on the relative importance of our predictive covariates
- Weights
  - Weights are chosen by minimizing the distance between observations $||X_1 - X_0 W||$ subject to some weight constraints
    - $W = (w_2, \dots, w_{J+1})'$ and $w_j \geq 0$ for $j = 2, \dots, J+1$
    - $w_2 + \cdots + w_{J+1} = 1$ such that there are no negative weights
- Importance Weights
  - Minimize synthetic control weights and Choice of V
    - Minimizing our distance between covariates requires
      - $\big||X_1 - X_0 W|\big| = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)}$
    - We will generate minimized weights V based upon a set of m covariates
    - $\sum_{m=1}^{k} V_m \left(X_{1m} - \sum_{j=2}^{J+1} w_j X_{jm}\right)^2$
    - Where v_m is the weight that reflects the importance that we assign to the mth variable when we compare the treated unit and the synthetic control unit
  - The choice of V is important because weights W depend on one's choice of V and synthetic control weights W*(V) is meant to reproduce the behavior of the outcome variable for the treated unit in the absence of treatment
- Choice of V
  - Most people choose a V that minimizes mean squared predicted error
    - $\sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^*(V) Y_{jt}\right)^2$
- Unobserved Factors
  - Abadie, et al (2010) argue, similar to parallel trends, only units that are alike on observed and unobserved factors would follow a similar trajectory in pre-treatment

Inference

- We need to calculate exact p-values from a distribution of parameters from placebo tests and compare it to the treatment parameter of interest
  - First, calculate placebo tests where each comparison donor gets the treatment
  - Second, calculate the Root Mean Squared Prediction Error in pre-treatment
  - Third, calculate the Root Mean Squared Prediction Error in post-treatment
  - Fourth, calculate the ratio of $\frac{RMSPE_{post}}{RMSPE_{pre}}$
  - Fifth, sort and rank the ratios in descending order from largest to smallest
  - Sixth, calculate the treatment units exact p-value in the distribution $p = \frac{Rank}{Total}$
  - You can plot the parameter of interest against the distribution from the placebo tests to look for outliers in pre-treatment compared to the treatment unit