Project 7

Diffe ences and Synthetic Control

Computational Social Science

1 Introduction

For this project, you will explore the question of whether the Affordable Care Act increased health insurance coverage (or conversely, decreased the number of people who are uninsured). The ACA was passed in March 2010, but several of its provisions were phased in over a few years. The ACA instituted the "individual mandate" which required that all Americans must carry health insurance, or else suffer a tax penalty. There are four mechanisms for how the ACA aims to reduce the uninsured population:

1. Require companies with more than 50 employees to provide health insurance.
2. Build state-run healthcare markets ("exchanges") for individuals to purchase health insurance.

3. Provide subsidies to middle income individuals and families who do not qualify for employer based coverage.

4. Expand Medicaid to require that states grant eligibility to all citizens and legal residents earning up to 138% of the federal poverty line. The federal government would initially pay 100% of the costs of this expansion, and over a period of 5 years the burden would shift so the federal government would pay 90% and the states would pay 10%.

In 2012, the Supreme Court heard the landmark case NFIB v. Sebelius, which principally challenged the constitutionality of the law under the theory that Congress could not institute an individual mandate. The Supreme Court ultimately upheld the individual mandate under Congress's taxation power, but struck down the requirement that states must expand Medicaid as impermissible subordination of the states to the federal government. Subsequently, several states refused to expand Medicaid when the program began on January 1, 2014. This refusal created the "Medicaid coverage gap" where there are individuals who earn too much to qualify for Medicaid under the old standards, but too little to qualify for the ACA subsidies targeted at middle-income individuals.

States that refused to expand Medicaid principally cited the cost as the primary [illegible] out however, that the decision not to expand primarily [illegible] isan lines. In the years since the initial expansion, seve[illegible] to the program, either because of a change in the gover[illegible] oters directly approved expansion via a ballot initiative.

You w[illegible] n of whether Medicaid expansion reduced the uninsured[illegible]. in the 7 years since it went into effect. To address t[illegible] e difference-in-differences estimation, and synthetic control.

2 Data

The dataset you will work with has been assembled from a few different sources about Medicaid. The key variables are:

- **State**: Full name of state
- **Medicaid Expansion Adoption**: Date that the state adopted the Medicaid expansion, if it did so.
- **Year**: Year of observation

- **Uninsured rate**: State uninsured rate in that year.

3 Exploratory Data Analysis

Create plots and provide 1-2 sentence analyses to answer the following questions:

1. Which states had the highest uninsured rates prior to 2014? The lowest?
2. Which states were home to most uninsured Americans prior to 2014? How about in the last year of the data available? **Note**: 2010 state population is provided as a variable to answer this question. In an actual study you would likely use population estimates over time, but to simplify you can assume these numbers stay about the same.

4 Difference-in-Differences Estimation

4.1 Estimate Model

Do the following:

1. Choose a state that adopted the Medicaid expansion on January 1, 2014 and a state that did not. **Hint**: Do not pick Massachusetts as it passed a universal healthcare law in 2006, and also avoid picking a state that adopted the Medicaid expansion between 2014 and 2015.

2. Ass[illegible] assumption for your choices using a plot. If you are [illegible] sumption has been met, pick another state and try [illegible] ates you tried).

3. Est[illegible] fferences estimate of the effect of the Medicaid exp[illegible] l share of the population. You may follow the lab[illegible] mate the differences in one pre-treatment and one[illegible], or take an average of the pre-treatment and pos[illegible]

4.2 Discussion Questions

1. Card/Krueger's original piece utilized the fact that towns on either side of the Delaware river are likely to be quite similar to one another in terms of demographics, economics, etc. Why is that intuition harder to replicate with this data?

2. What are the strengths and weaknesses of using the parallel trends assumption in difference-in-differences estimates?

5 Synthetic Control

5.1 Estimate Synthetic Control

Although several states did not expand Medicaid on January 1, 2014, many did later on. In some cases, a Democratic governor was elected and pushed for a state budget that included the Medicaid expansion, whereas in others voters approved expansion via a ballot initiative. The 2018 election was a watershed moment where several Republican-leaning states elected Democratic governors and approved Medicaid expansion. In cases with a ballot initiative, the state legislature and governor still must implement the results via legislation. For instance, Idaho voters approved a Medicaid expansion in the 2018 election, but it was not implemented in the state budget until late 2019, with enrollment beginning in 2020.

Do the following:

1. Choose a state that adopted the Medicaid expansion *after* January 1, 2014. Construct a non-augmented synthetic control and plot the results (both pre-treatment fit and post-treatment differences). Also report the average ATT and L2 imbalance.

2. Re-run the same analysis but this time use an augmentation (default choices are Ridge, Matrix Completion, and GSynth). Create the same plot and report the average ATT and L2 imbalance.

3. Plot barplots to visualize the weights of the donors.

HINT: Is there any preprocessing you need to do before you allow the program to find weights for donor states?

5.2 Discussions

1. What are the advantages and disadvantages of synthetic control compared to other estimators?
2. One limitation of synthetic control is that the weights are bounded between [0,1] and the weights must sum to 1. Augmentation might relax this assumption by allowing for negative weights. Does this create an interpretation problem, and how should we balance this consideration against the improvements augmentation offers in terms of imbalance in the pre-treatment period?

6 Staggered Adoption Synthetic Control

6.1 Estimate Multisynth

Do the following:

1. Estimate a multisynth model that treats each state individually. Choose a fraction of states that you can fit on a plot and examine their treatment effects.

2. Estimate a multisynth model using time cohorts. For the purpose of this exercise, you can simplify the treatment time so that states that adopted Medicaid expansion within the same year (i.e. all states that adopted epxansion in 2016) count for the same cohort. Plot the treatment effects for these time cohorts.

6.2 Discussion Questions

1. One feature of Medicaid is that it is jointly administered by the federal government and the states, and states have some flexibility in how they implement Medicaid. For example, during the Trump administration, several states applied for waivers where they could add work requirements to the eligibility standards (i.e. an individual needed to work for 80 hours/month to qualify for Medicaid). Given these differences, do you see evidence for the idea that different states had different treatment effect sizes?
2. Do you see evidence for the idea that early adopters of Medicaid expansion enjoyed a larger decrease in the uninsured population?

7 General Discussion Questions

1. Wh ic control estimates well suited to studies of agg states, countries, etc?
2. Wh o treatment play in DiD/synthetic control versus ? When would we want to use either method?