Project 7: Difference-in-Differences and Synthetic Control

```r
# Install and load packages
if (!require("pacman")) install.packages("pacman")
```

```
## Loading required package: pacman
```

```r
devtools::install_github("ebenmichael/augsynth")
```

```
## Using GitHub PAT from the git credential store.
```

```
## Skipping install of 'augsynth' from a github remote, the SHA1 (0f4f1bcc) has not changed since last i
##   Use `force = TRUE` to force installation
```

```r
#install.packages("usmap")
library(usmap)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.0     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(ggthemes)
library(augsynth)
library(gsynth)
```

```
## ## Syntax has been updated since v.1.2.0.
## ## Comments and suggestions -> yiqingxu@stanford.edu.
```

```r
# set seed
set.seed(44)

# load data
# Kasey, if you have to run this, I changed the path to match where the data is relative to where the f
medicaid_expansion <- read_csv('~/git/Computational-Social-Science-Training-Program/Projects/Project 7/d
```

```
## Rows: 663 Columns: 5
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (1): State
## dbl  (3): year, uninsured_rate, population
## date (1): Date_Adopted
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

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:

- Require companies with more than 50 employees to provide health insurance.
- Build state-run healthcare markets ("exchanges") for individuals to purchase health insurance.
- Provide subsidies to middle income individuals and families who do not qualify for employer based coverage.
- 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 indivudals 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 factor. Critics pointed out however, that the decision not to expand primarily broke down along partisan lines. In the years since the initial expansion, several states have opted into the program, either because of a change in the governing party, or because voters directly approved expansion via a ballot initiative.

You will explore the question of whether Medicaid expansion reduced the uninsured population in the U.S. in the 7 years since it went into effect. To address this question, you will use difference-in-differences estimation, and synthetic control.

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.

Exploratory Data Analysis

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

- Which states had the highest uninsured rates prior to 2014? The lowest?
- **Answer**: Nevada has the highest average uninsured rate, Massachusetts has the lowest average uninsured rate before 2014.
- Which states were home to most uninsured Americans prior to 2014? How about in the last year in the data set? **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.
- **Answer**: Prior to 2014, California had the most uninsured Americans, followed by Texas. In 2020, Texas had the most uninsured Americans, followed by California.

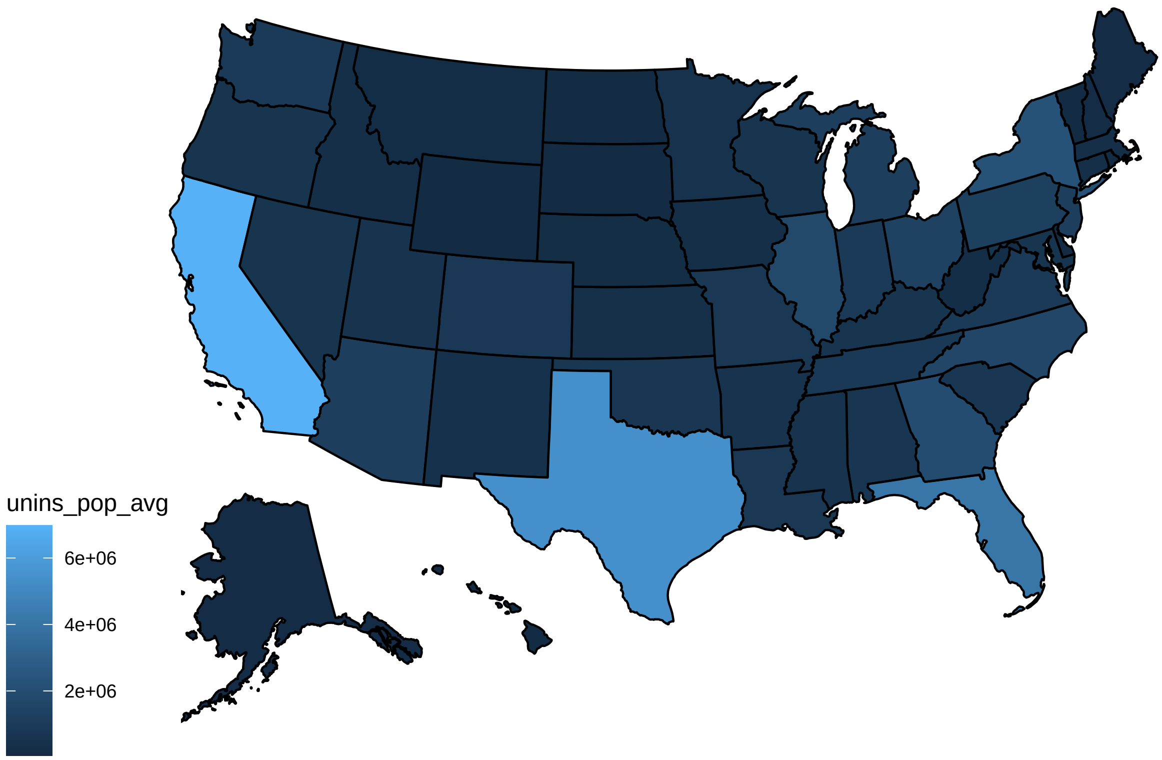
```r
# highest and lowest uninsured rates

med_sub <- medicaid_expansion %>%
  filter(year<2014) %>%
  group_by(State) %>%
  summarize(unins_avg = mean(uninsured_rate))

med_sub$State[which.max(med_sub$unins_avg)] #Nevada is highest uninsured rate
```

```
## [1] "Nevada"
```

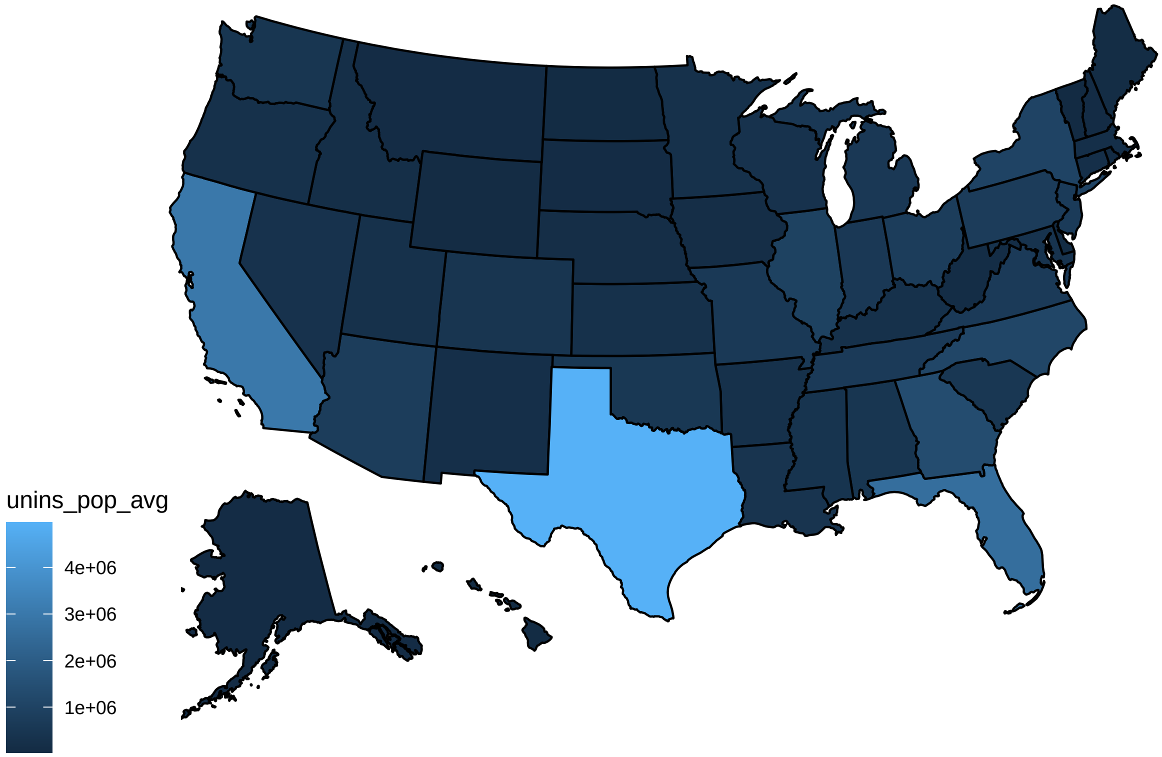
```r
med_sub$State[which.min(med_sub$unins_avg)] #Massachusetts is lowest uninsured rate
```

```
## [1] "Massachusetts"
```

```r
# most uninsured Americans
q2 <- medicaid_expansion %>%
  filter(year < 2014) %>%
  mutate(unins_pop = uninsured_rate*population) %>%
  group_by(State) %>%
  summarize(unins_pop_avg = mean(unins_pop))

q2 <- q2 %>%
  rename(state = State)

#visualization prior to 2014
plot_usmap(data = q2, values = "unins_pop_avg")
```




```r
# 2020 most uninsured Americans
q2_2020 <- medicaid_expansion %>%
  filter(year == 2020) %>%
   mutate(unins_pop = uninsured_rate*population) %>%
  group_by(State) %>%
  summarize(unins_pop_avg = mean(unins_pop))

q2_2020 <- q2_2020 %>%
  rename(state = State)

plot_usmap(data = q2_2020, values = "unins_pop_avg")
```



Difference-in-Differences Estimation

Estimate Model

Do the following:

- 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.
- Assess the parallel trends assumption for your choices using a plot. If you are not satisfied that the assumption has been met, pick another state and try again (but detail the states you tried). **Answer**: I ultimately selected Iowa and Georgia, which had essentially parallel trends before 2014. There was a slight difference in the slopes from 2010-2012, but the shapes look very similar broadly.

```
# Parallel Trends plot

#Selecting Arizona as 2014 adopter
medicaid_expansion %>%
 filter(Date_Adopted == "2014-01-01")
```

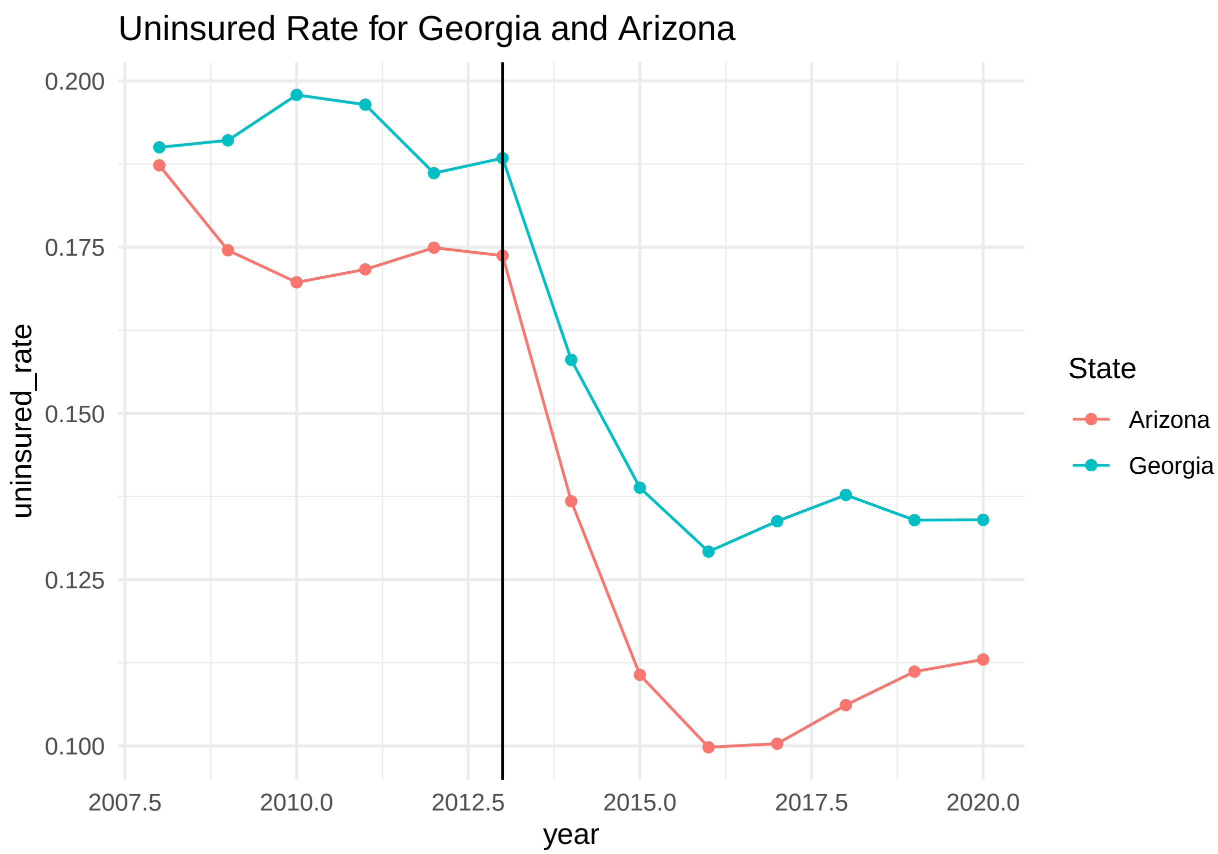
```
## # A tibble: 325 x 5
##    State                  Date_Adopted  year uninsured_rate population
##    <chr>                  <date>       <dbl>          <dbl>      <dbl>
```

```
##  1 Arizona              2014-01-01    2008         0.187     6731484
##  2 Arkansas             2014-01-01    2008         0.179     2994079
##  3 California           2014-01-01    2008         0.178    38802500
##  4 Colorado             2014-01-01    2008         0.170     5355856
##  5 Connecticut          2014-01-01    2008         0.0891    3596677
##  6 Delaware             2014-01-01    2008         0.108      935614
##  7 District of Columbia 2014-01-01    2008         0.0805         NA
##  8 Hawaii               2014-01-01    2008         0.0672    1419561
##  9 Illinois             2014-01-01    2008         0.129    12880580
## 10 Iowa                 2014-01-01    2008         0.0873    3107126
## # i 315 more rows
```

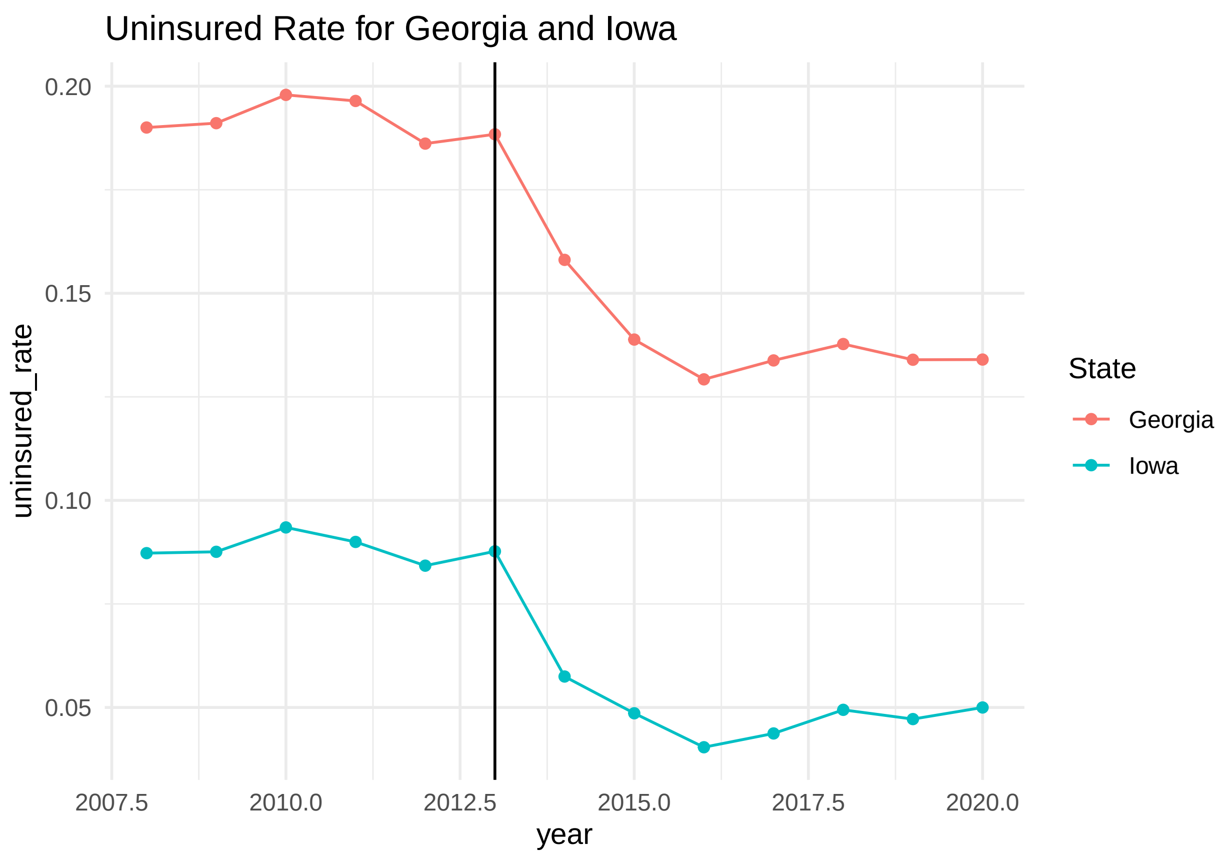
```r
#selecting Georgia as non-adopter - was pretty sure GA hasn't adopted but wanted to check
medicaid_expansion %>%
 filter(State=="Georgia")
```

```
## # A tibble: 13 x 5
##    State   Date_Adopted  year uninsured_rate population
##    <chr>   <date>       <dbl>          <dbl>      <dbl>
##  1 Georgia NA            2008          0.190   10097343
##  2 Georgia NA            2009          0.191   10097343
##  3 Georgia NA            2010          0.198   10097343
##  4 Georgia NA            2011          0.196   10097343
##  5 Georgia NA            2012          0.186   10097343
##  6 Georgia NA            2013          0.188   10097343
##  7 Georgia NA            2014          0.158   10097343
##  8 Georgia NA            2015          0.139   10097343
##  9 Georgia NA            2016          0.129   10097343
## 10 Georgia NA            2017          0.134   10097343
## 11 Georgia NA            2018          0.138   10097343
## 12 Georgia NA            2019          0.134   10097343
## 13 Georgia NA            2020          0.134   10097343
```

```r
# parallel trends
medicaid_expansion %>%
  filter(State %in% c("Arizona", "Georgia")) %>%
  ggplot() +
  geom_point(aes(x=year,
                 y=uninsured_rate,
                 color = State)) +
  geom_line(aes(x=year,
                 y=uninsured_rate,
                 color = State)) +
    geom_vline(aes(xintercept = 2013)) +
  ggtitle('Uninsured Rate for Georgia and Arizona') +
  theme_minimal()
```




```r
#Picking new state to match GA better
medicaid_expansion %>%
  filter(State %in% c("Iowa", "Georgia")) %>%
  ggplot() +
  geom_point(aes(x=year,
                 y=uninsured_rate,
                 color = State)) +
  geom_line(aes(x=year,
                 y=uninsured_rate,
                 color = State)) +
    geom_vline(aes(xintercept = 2013)) +
  ggtitle('Uninsured Rate for Georgia and Iowa') +
  theme_minimal()
```



- Estimates a difference-in-differences estimate of the effect of the Medicaid expansion on the uninsured share of the population. You may follow the lab example where we estimate the differences in one pre-treatment and one post-treatment period, or take an average of the pre-treatment and post-treatment outcomes

```
# Difference-in-Differences estimation

#dataset with just GA and IA
gi <- medicaid_expansion %>%
  filter(State %in% c("Iowa", "Georgia"))

# pre-treatment difference
pre_treat <- gi %>%
  filter(year <= 2013) %>%
  select(State, uninsured_rate) %>%
  group_by(State) %>%
  summarize(avg = mean(uninsured_rate)) %>%
  pivot_wider(names_from = State,
              values_from = avg) %>%
  summarize(Iowa - Georgia)

# post-treatment difference
post_treat <- gi %>%
  filter(year > 2013) %>%
  select(State, uninsured_rate) %>%
```

```
  group_by(State) %>%
  summarize(avg = mean(uninsured_rate)) %>%
  pivot_wider(names_from = State,
              values_from = avg) %>%
  summarize(Iowa - Georgia)

# diff in diff

diff_in_diff <- post_treat - pre_treat
diff_in_diff
```

```
##   Iowa - Georgia
## 1     0.01346215
```

Answer: In the case of Georgia and Iowa, we see a treatment effect of 0.0135 on the percent of the population uninsured.

Discussion Questions

- 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?
- **Answer**: In this data, it's hard to say that two states - even ones that neighbor one another - are as similar as the towns in NJ/PA. With a larger land mass, most states have different economic, cultural, demographic, etc. factors. Even though we assessed parallel trends, I can't say that Georgia and Iowa are similar in the ways that the two towns in the Card/Kreuger piece were.
- What are the strengths and weaknesses of using the parallel trends assumption in difference-in-differences estimates?
- **Answer**: Strengths of the parallel trends assumption are that it attempts to be a proxy for the factors that influence the outcome - if the outcome variable has parallel trends in both cases, we can make an assumption that changes in the outcome post-treatment are due to the treatment alone. It is also computationally efficient compared to something like regression discontinuity, and allows for causal inference. However, the weakness of this assumption is that it is easy to violate, and difficult to prove that various demographic factors are truly similar enough (particularly in the case of US states) to make the cases comparable.

Synthetic Control

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:

- 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.

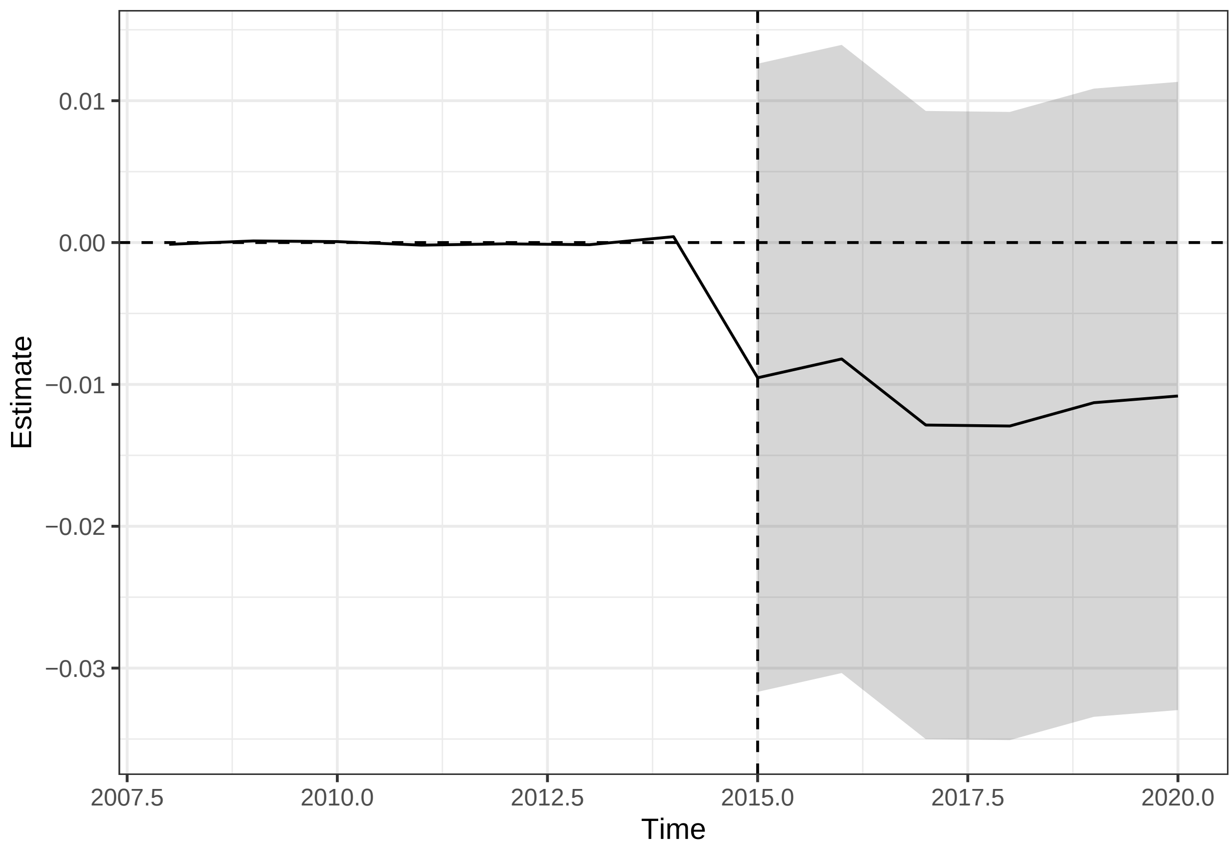
```r
# non-augmented synthetic control
# selecting Pennsylvania, adopted Jan 1 2015

pa <- medicaid_expansion %>%
  mutate(treatment = case_when(State == "Pennsylvania" & year >=2015 ~ 1,
                               TRUE ~ 0))

syn <-                              # save object
  augsynth(uninsured_rate ~ treatment,
                         State,      # unit
                         year,  # time
                         pa,    # data
           progfunc = "None",        # plain syn control
           scm = T)                  # synthetic control
```

```
## One outcome and one treatment time found. Running single_augsynth.
```

```r
# Avg ATT = -0.0109, L2 imbalance = 0.001
summary(syn)
```

```
##
## Call:
## single_augsynth(form = form, unit = !!enquo(unit), time = !!enquo(time),
##     t_int = t_int, data = data, progfunc = "None", scm = ..2)
##
## Average ATT Estimate (p Value for Joint Null):  -0.0109   ( 0.61 )
## L2 Imbalance: 0.001
## Percent improvement from uniform weights: 99.5%
##
## Avg Estimated Bias: NA
##
## Inference type: Conformal inference
##
##  Time Estimate 95% CI Lower Bound 95% CI Upper Bound p Value
##  2015   -0.010             -0.032              0.013   0.228
##  2016   -0.008             -0.030              0.014   0.258
##  2017   -0.013             -0.035              0.009   0.255
##  2018   -0.013             -0.035              0.009   0.365
##  2019   -0.011             -0.033              0.011   0.256
##  2020   -0.011             -0.033              0.011   0.237
```

```r
plot(syn)
```



- 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.

```
# augmented synthetic control
# selecting Pennsylvania, adopted Jan 1 2015

aug_syn <-                              # save object
  augsynth(uninsured_rate ~ treatment,
                         State,     # unit
                         year,  # time
                         pa,    # data
           progfunc = "ridge",  # specify ridge augmentation
           scm = T)             # synthetic control
```

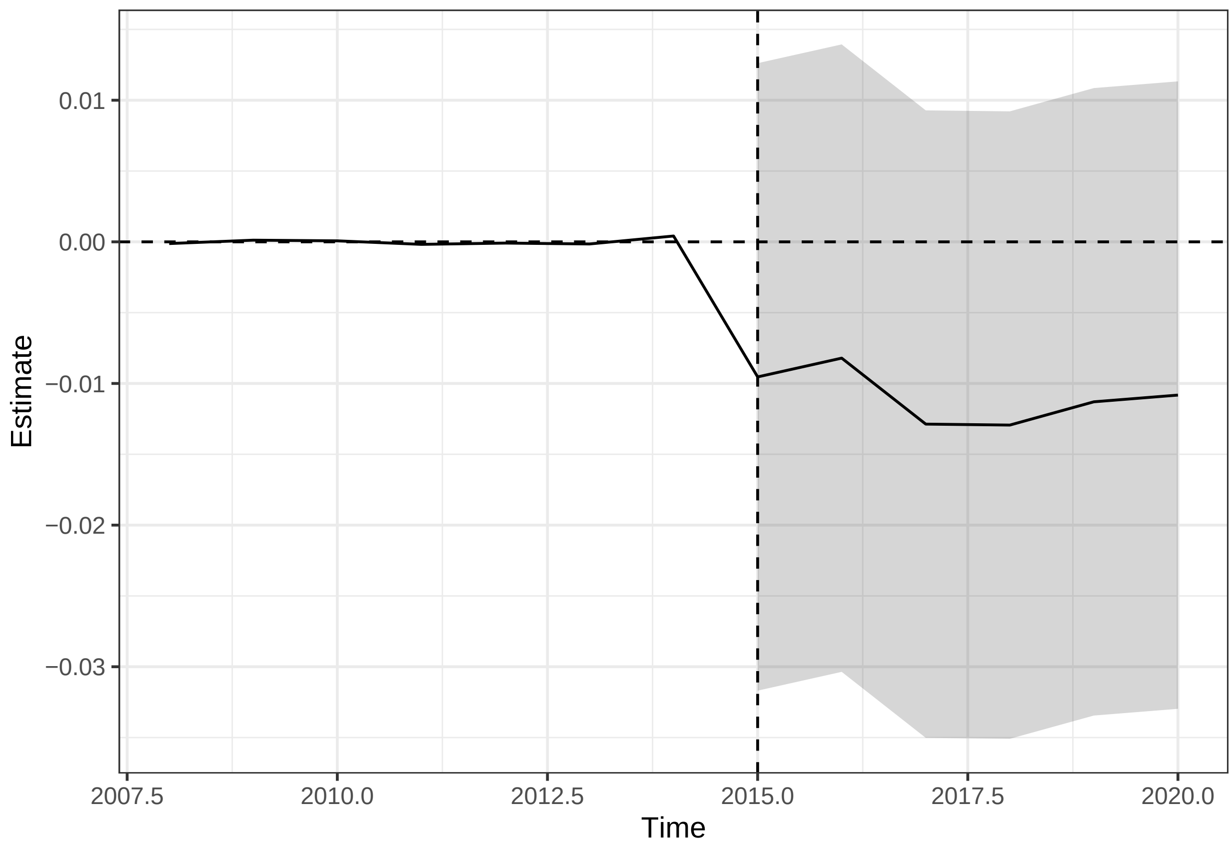
```
## One outcome and one treatment time found. Running single_augsynth.
```

```
# Avg ATT = -0.0109, L2 imbalance = 0.001
summary(aug_syn)
```

```
##
## Call:
## single_augsynth(form = form, unit = !!enquo(unit), time = !!enquo(time),
##     t_int = t_int, data = data, progfunc = "ridge", scm = ..2)
##
```

```
## Average ATT Estimate (p Value for Joint Null):  -0.0109   ( 0.61 )
## L2 Imbalance: 0.001
## Percent improvement from uniform weights: 99.5%
##
## Avg Estimated Bias: 0.000
##
## Inference type: Conformal inference
##
##  Time Estimate 95% CI Lower Bound 95% CI Upper Bound p Value
##  2015   -0.010             -0.032              0.013   0.257
##  2016   -0.008             -0.030              0.014   0.253
##  2017   -0.013             -0.035              0.009   0.253
##  2018   -0.013             -0.035              0.009   0.370
##  2019   -0.011             -0.033              0.011   0.252
##  2020   -0.011             -0.033              0.011   0.245
```

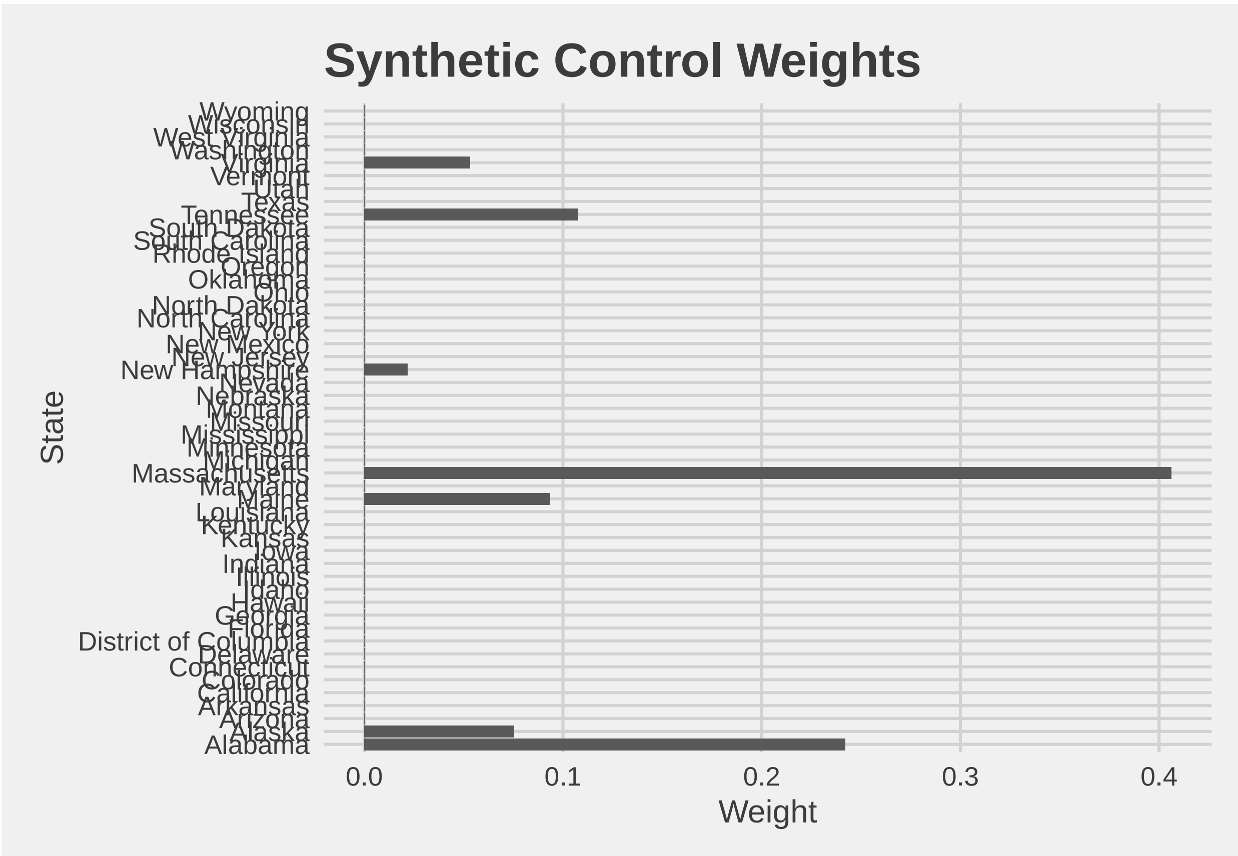
```
plot(aug_syn)
```



- Plot barplots to visualize the weights of the donors.

```
# barplots of weights

data.frame(aug_syn$weights) %>%
  tibble::rownames_to_column('State') %>%
  ggplot() +
```

```r
  geom_bar(aes(x = State, y = aug_syn.weights),
           stat = 'identity') +
  coord_flip() + # coord flip
  theme_fivethirtyeight() +
  theme(axis.title = element_text()) +
  ggtitle('Synthetic Control Weights') +
  xlab('State') +
  ylab('Weight')
```



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

Discussion Questions

- What are the advantages and disadvantages of synthetic control compared to difference-in-differences estimators?
- **Answer**: Advantages of synthetic control over diff-in-diff is that it avoids the problem of selecting a comparison case, using a weighted average of all non-treated cases to a synthetic match of the treated group. This allows for a "direct" comparison to the synthetic version compared to looking at differences between treatment and control before and after treatment. The primary disadvantage is that synthetic control is harder to interpret - what really is synthetic Pennsylvania, and how do we know it's the "true" counterfactual? - and the there are strong validity assumptions required to use the weights.
- One of the benefits 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?

- **Answer**: Negative weights would definitely make interpretation more difficult, but might be useful for improving the pre-treatment balance with the treated unit. Ways to balance this interpretation problem would be to conduct sensitivity analyses comparing augmentation with negative weights versus only positive ones, or use qualitative information to motivate the negative weights.

Staggered Adoption Synthetic Control

Estimate Multisynth

Do the following:

- 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.

```
# multisynth model states

#selecting southeastern states: Florida, Georgia, Kentucky, North Carolina, South Carolina, Tennessee,
multisynth <- medicaid_expansion %>%
  filter(State %in% c("Florida", "Georgia", "Kentucky", "North Carolina", "South Carolina", "Tennessee"
  mutate(year_adopted = as.numeric(format(Date_Adopted,'%Y'))) %>%
  mutate(treated = ifelse(year >= year_adopted, 1, 0))


ppool_syn <- multisynth(uninsured_rate ~ treated,
                        State,                        # unit
                        year,                         # time
                        multisynth,  # data
                        nu = 0,                        # treat each state individually
                        n_leads = 10)                 # post-treatment periods to estimate
print(ppool_syn$nu)
```

```
## [1] 0
```

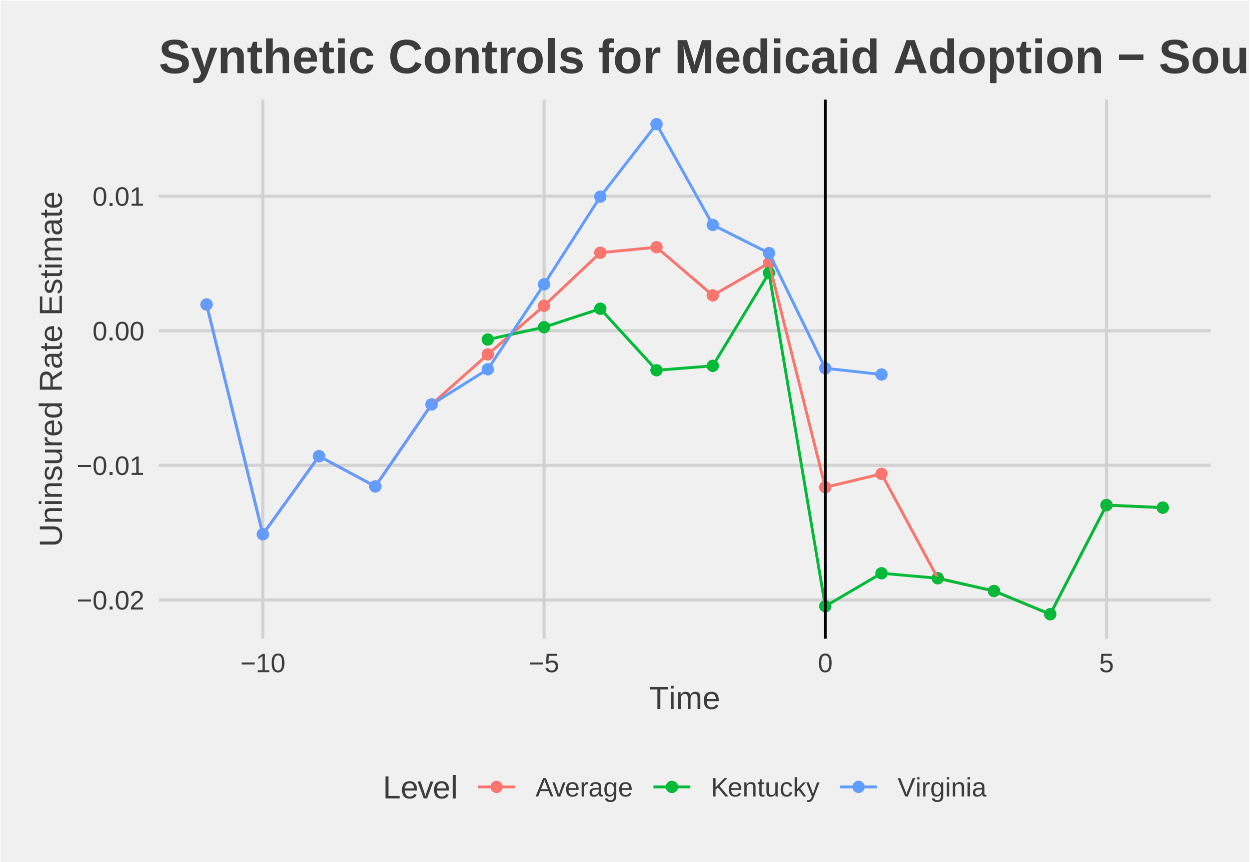
```
ppool_syn
```

```
##
## Call:
## multisynth(form = uninsured_rate ~ treated, unit = State, time = year,
##     data = multisynth, n_leads = 10, nu = 0)
##
## Average ATT Estimate: -0.010
```

```
ppool_syn_summ <- summary(ppool_syn)

ppool_syn_summ$att %>%
  ggplot(aes(x = Time, y = Estimate, color = Level)) +
  geom_point() +
```

```
  geom_line() +
  geom_vline(xintercept = 0) +
  theme_fivethirtyeight() +
  theme(axis.title = element_text(),
        legend.position = "bottom") +
  ggtitle('Synthetic Controls for Medicaid Adoption - Southeastern States') +
  xlab('Time') +
  ylab('Uninsured Rate Estimate')
```

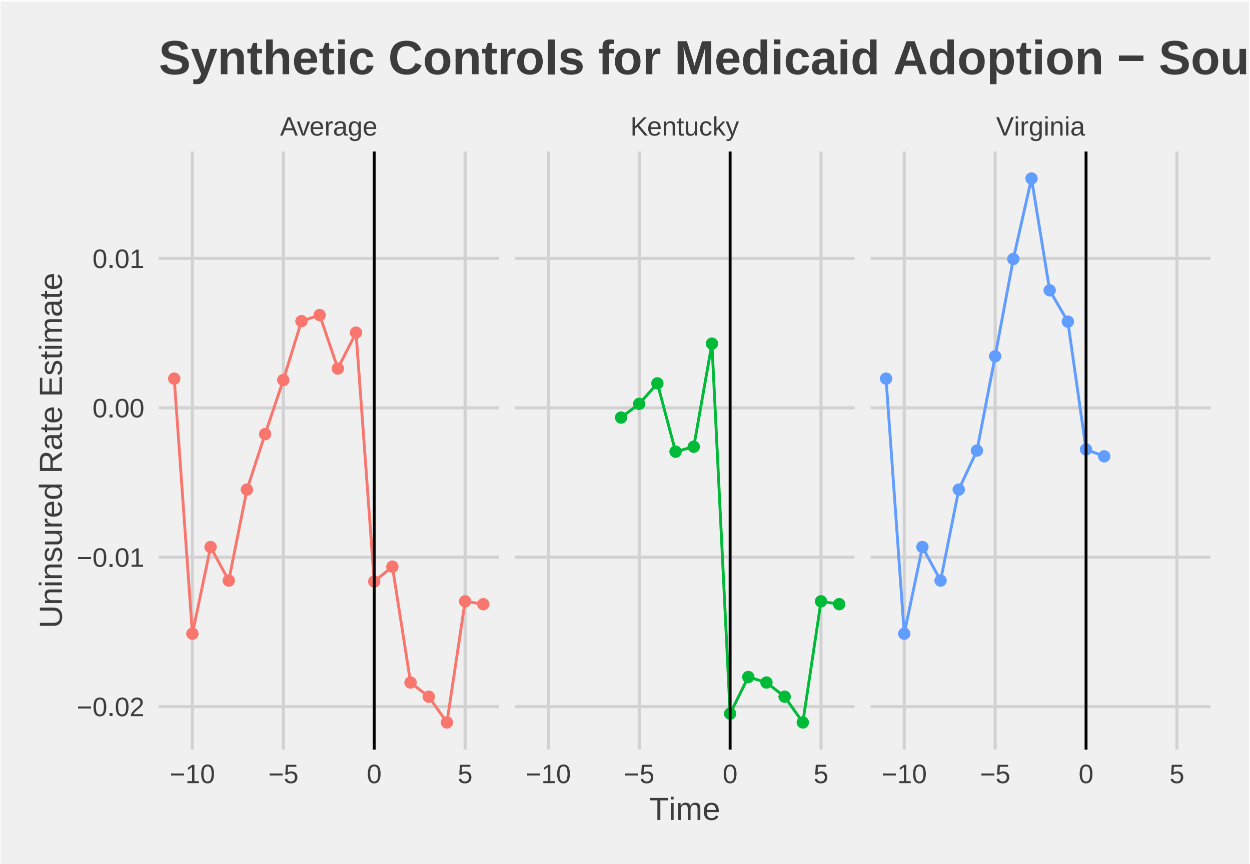
```
## Warning: Removed 13 rows containing missing values or values outside the scale range
## (‘geom_point()‘).

## Warning: Removed 13 rows containing missing values or values outside the scale range
## (‘geom_line()‘).
```




```
ppool_syn_summ$att %>%
  ggplot(aes(x = Time, y = Estimate, color = Level)) +
  geom_point() +
  geom_line() +
  geom_vline(xintercept = 0) +
  theme_fivethirtyeight() +
  theme(axis.title = element_text(),
        legend.position = 'None') +
  ggtitle('Synthetic Controls for Medicaid Adoption - Southeastern States') +
```

```
  xlab('Time') +
  ylab('Uninsured Rate Estimate') +
  facet_wrap(~Level)
```

```
## Warning: Removed 13 rows containing missing values or values outside the scale range
## (`geom_point()`).
## Removed 13 rows containing missing values or values outside the scale range
## (`geom_line()`).
```



- 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.

```
# multisynth model time cohorts
#already simplified treatment time

ppool_syn_time <- multisynth(uninsured_rate ~ treated,
                             State,                          # unit
                             year,                           # time
                             multisynth,  # data
                             nu = 0,                          # treat each state individually
                             n_leads = 10,       # post-treatment periods to estimate
                             time_cohort = TRUE)
print(ppool_syn_time$nu)
```

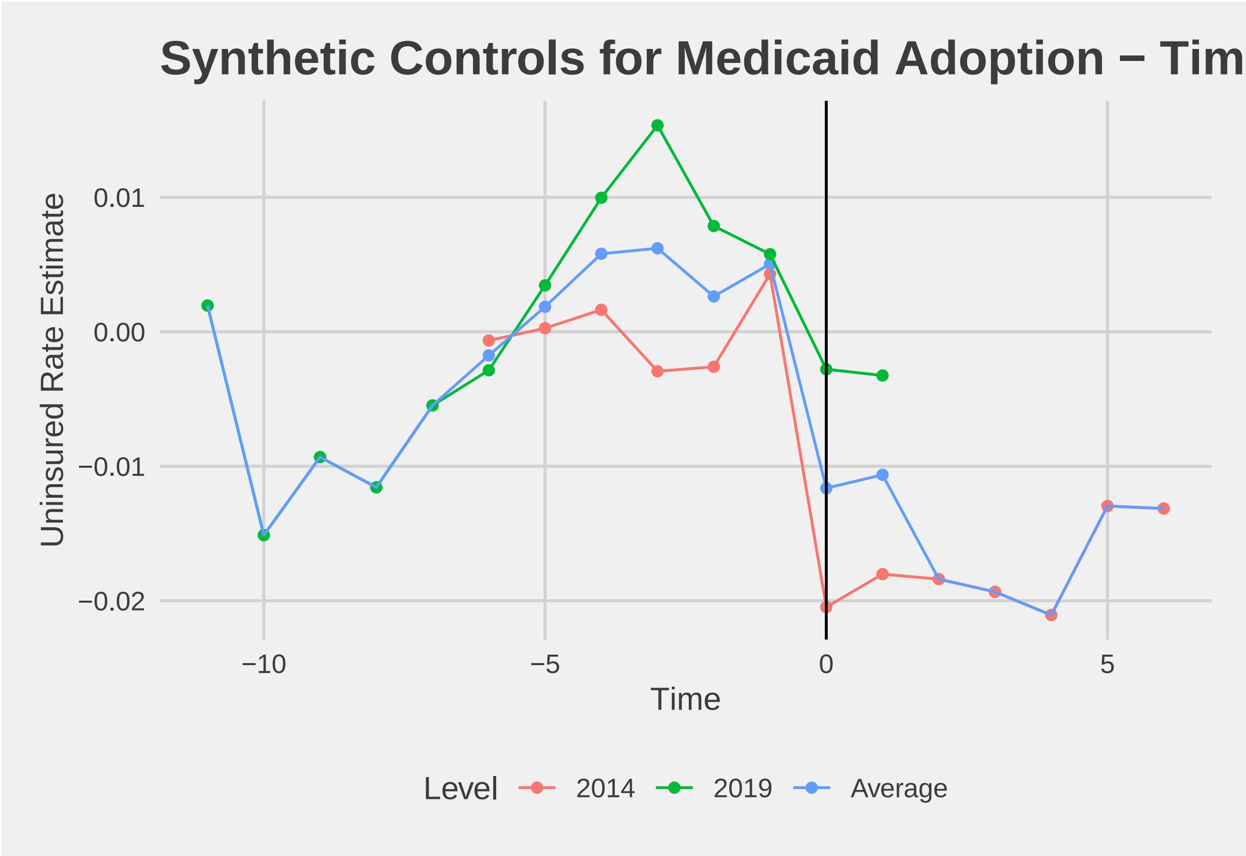
```
## [1] 0
```

```r
ppool_syn_time
```

```
##
## Call:
## multisynth(form = uninsured_rate ~ treated, unit = State, time = year,
##     data = multisynth, n_leads = 10, nu = 0, time_cohort = TRUE)
##
## Average ATT Estimate: -0.010
```

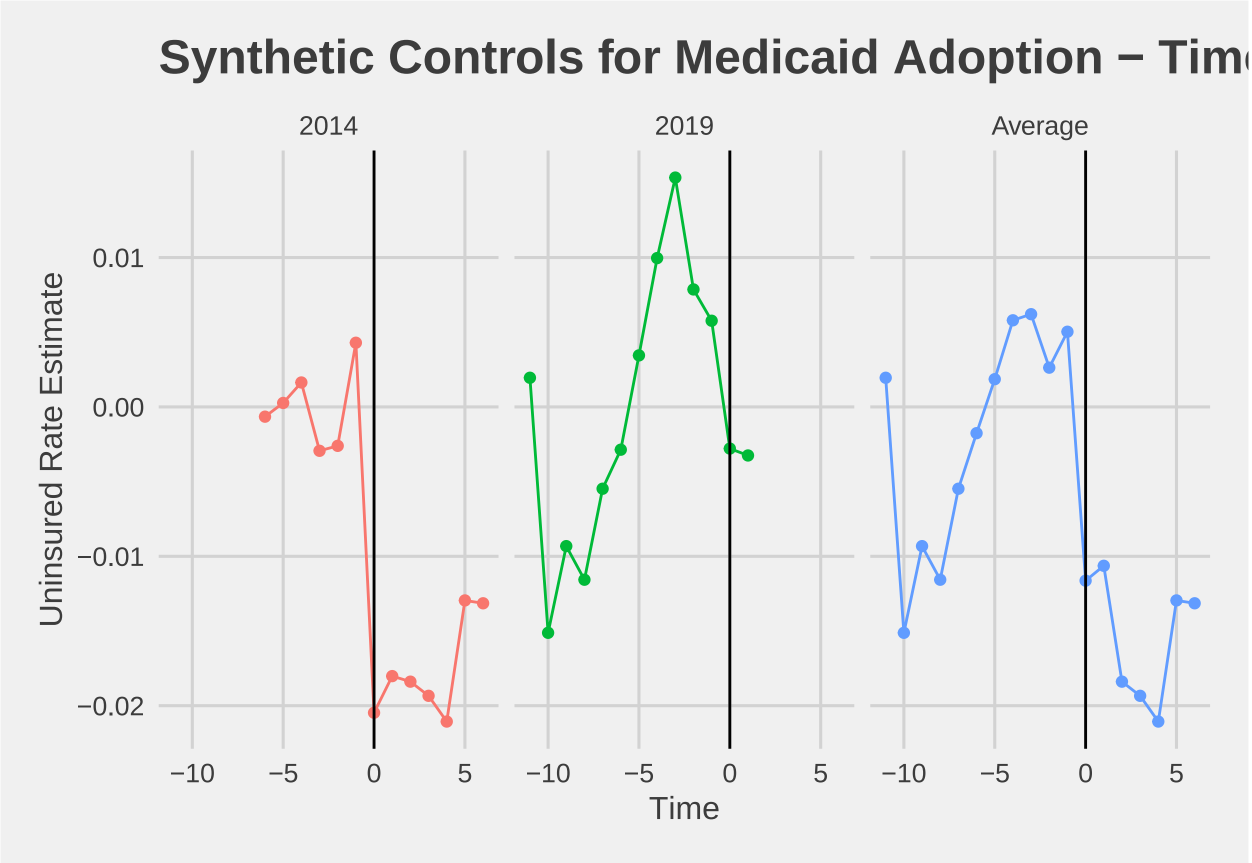
```r
ppool_syn_time_summ <- summary(ppool_syn_time)

ppool_syn_time_summ$att %>%
  ggplot(aes(x = Time, y = Estimate, color = Level)) +
  geom_point() +
  geom_line() +
  geom_vline(xintercept = 0) +
  theme_fivethirtyeight() +
  theme(axis.title = element_text(),
        legend.position = "bottom") +
  ggtitle('Synthetic Controls for Medicaid Adoption - Time Cohorts + Southeastern States') +
  xlab('Time') +
  ylab('Uninsured Rate Estimate')
```

```
## Warning: Removed 13 rows containing missing values or values outside the scale range
## (`geom_point()`).

## Warning: Removed 13 rows containing missing values or values outside the scale range
## (`geom_line()`).
```




```r
ppool_syn_time_summ$att %>%
  ggplot(aes(x = Time, y = Estimate, color = Level)) +
  geom_point() +
  geom_line() +
  geom_vline(xintercept = 0) +
  theme_fivethirtyeight() +
  theme(axis.title = element_text(),
        legend.position = 'None') +
  ggtitle('Synthetic Controls for Medicaid Adoption - Time Cohorts + Southeastern States') +
  xlab('Time') +
  ylab('Uninsured Rate Estimate') +
  facet_wrap(~Level)
```

```
## Warning: Removed 13 rows containing missing values or values outside the scale range
## (`geom_point()`).
## Removed 13 rows containing missing values or values outside the scale range
## (`geom_line()`).
```



Discussion Questions

- 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?
- **Answer**: Yes, there is definitely evidence for different treatment effect sizes for a variety of reasons. In addition to adminstrative differences at the state level, cultural and demographic differences could influence uptake of Medicaid even after a state participates in the expansion. For example, perhaps a state with a larger rural population and fewer healthcare centers would see lower uptake after passing an expansion, particularly after the penalty was eliminated nationwide.
- Do you see evidence for the idea that early adopters of Medicaid expansion enjoyed a larger decrease in the uninsured population?
- **Answer**: Based on the time cohort graphs, it does seem like early adopters of the expansion experienced a larger decrease in the uninsured population, but it's hard to tell since there weren't any adopters in 2017 or 2018. However, given the political attitudes toward the ACA at the time, it's possible that this difference is driven by other factors besides year of adoption.

General Discussion Questions

- Why are DiD and synthetic control estimates well suited to studies of aggregated units like cities, states, countries, etc?
- **Answer**: Both DiD and synthetic control estimates attempt to address (in different ways) unobserved heterogeneity across cities, states, countries, etc. that influences the outcome and the treatment - policies are not passed in a vacuum. In individual-level studies, there are other strategies to account for this like random assignment.
- What role does selection into treatment play in DiD/synthetic control versus regression discontinuity? When would we want to use either method?
- **Answer**: DiD and synthetic controls adjust for selection into treatment through the assumption that the outcome would have had a similar trend as the control group had the treatment not been applied. RDD addresses selection into treatment by using the fact that treatment is assigned based on a continuous variable that is related to the outcome of interests. This allows for causal claims by comparing the units just above and just below the cutoff threshold.

We want to use DiD or synthetic controls when the treatment is assigned at a group level (like states, etc.), and there are clear control units. RDD is a good fit where the treatment (assignment variable) is continuous, there's a clear threshold for treatment assignment (i.e. elections, rankings, etc), and there are no meaningful differences in units right below or above the threshold.