Statistical Report on Qualified Immunity

Setup

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.0     v dplyr   1.0.5
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

##
## Attaching package: 'rvest'

## The following object is masked from 'package:readr':
##
##     guess_encoding

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union

## Loading required package: NLP

##
## Attaching package: 'NLP'

## The following object is masked from 'package:ggplot2':
##
##     annotate

## ##
## ## Synth Package: Implements Synthetic Control Methods.

## ## See http://www.mit.edu/~jhainm/software.htm for additional information.
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   INCIDENT_ID = col_double(),
##   OFFENSE_ID = col_double(),
##   OFFENSE_CODE = col_character(),
##   OFFENSE_CODE_EXTENSION = col_double(),
##   OFFENSE_TYPE_ID = col_character(),
##   OFFENSE_CATEGORY_ID = col_character(),
##   FIRST_OCCURRENCE_DATE = col_character(),
##   LAST_OCCURRENCE_DATE = col_character(),
##   REPORTED_DATE = col_character(),
##   INCIDENT_ADDRESS = col_character(),
##   GEO_X = col_double(),
##   GEO_Y = col_double(),
##   GEO_LON = col_double(),
##   GEO_LAT = col_double(),
##   DISTRICT_ID = col_double(),
##   PRECINCT_ID = col_double(),
##   NEIGHBORHOOD_ID = col_character(),
##   IS_CRIME = col_double(),
##   IS_TRAFFIC = col_double()
## )

##
## -- Column specification --------------------------------------------------------
## cols(
##   `Report Number` = col_character(),
##   `Offense ID` = col_double(),
##   `Offense Start DateTime` = col_character(),
##   `Offense End DateTime` = col_character(),
##   `Report DateTime` = col_character(),
##   `Group A B` = col_character(),
##   `Crime Against Category` = col_character(),
##   `Offense Parent Group` = col_character(),
##   Offense = col_character(),
##   `Offense Code` = col_character(),
##   Precinct = col_character(),
##   Sector = col_character(),
##   Beat = col_character(),
##   MCPP = col_character(),
##   `100 Block Address` = col_character(),
##   Longitude = col_double(),
##   Latitude = col_double()
## )

##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_character(),
##   X = col_double(),
##   Y = col_double(),
##   precinct = col_double(),
##   reportedTime = col_double(),
```

```
##   beginTime = col_double(),
##   centergbsid = col_double(),
##   centerLong = col_double(),
##   centerLat = col_double(),
##   centerX = col_double(),
##   centerY = col_double(),
##   OBJECTID = col_double()
## )
## i Use `spec()` for the full column specifications.

## Warning: 3 parsing failures.
##  row      col expected actual                                  file
## 1573 precinct a double     UI 'data/Police_Incidents_2018_PIMS.csv'
## 8438 precinct a double     UI 'data/Police_Incidents_2018_PIMS.csv'
## 8454 precinct a double     UI 'data/Police_Incidents_2018_PIMS.csv'

##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_character(),
##   X = col_double(),
##   Y = col_double(),
##   reportedTime = col_double(),
##   beginTime = col_double(),
##   centergbsid = col_double(),
##   centerLong = col_double(),
##   centerLat = col_double(),
##   centerX = col_double(),
##   centerY = col_double(),
##   OBJECTID = col_double()
## )
## i Use `spec()` for the full column specifications.

##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_character(),
##   X = col_double(),
##   Y = col_double(),
##   reportedTime = col_double(),
##   beginTime = col_double(),
##   centergbsid = col_double(),
##   centerLong = col_double(),
##   centerLat = col_double(),
##   centerX = col_double(),
##   centerY = col_double(),
##   OBJECTID = col_double()
## )
## i Use `spec()` for the full column specifications.

##
## -- Column specification --------------------------------------------------------
## cols(
```

```
##   .default = col_character(),
##   X = col_double(),
##   Y = col_double(),
##   precinct = col_double(),
##   reportedTime = col_double(),
##   beginTime = col_double(),
##   centergbsid = col_double(),
##   centerLong = col_double(),
##   centerLat = col_double(),
##   centerX = col_double(),
##   centerY = col_double(),
##   OBJECTID = col_double()
## )
## i Use `spec()` for the full column specifications.
```

```
## Warning: 3 parsing failures.
##  row      col expected actual                              file
## 4338 precinct a double     UI 'data/Police_Incidents_2021.csv'
## 4590 precinct a double     UI 'data/Police_Incidents_2021.csv'
## 5945 precinct a double     UI 'data/Police_Incidents_2021.csv'
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   Incident = col_double(),
##   `Occurrence
## Date` = col_character(),
##   `Occurrence
## Hour` = col_double(),
##   `NIBRS
## Class` = col_character(),
##   NIBRSDescription = col_character(),
##   `Offense
## Count` = col_double(),
##   Beat = col_character(),
##   Premise = col_character(),
##   `Block Range` = col_character(),
##   StreetName = col_character(),
##   `Street
## Type` = col_character(),
##   Suffix = col_character(),
##   City = col_character(),
##   `ZIP Code` = col_character()
## )
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   Incident = col_double(),
##   `Occurrence
## Date` = col_character(),
##   `Occurrence
## Hour` = col_double(),
```

```
##   ‘NIBRS
## Class‘ = col_character(),
##   NIBRSDescription = col_character(),
##   ‘Offense
## Count‘ = col_double(),
##   Beat = col_character(),
##   Premise = col_character(),
##   ‘Block Range‘ = col_character(),
##   StreetName = col_character(),
##   ‘Street
## Type‘ = col_character(),
##   Suffix = col_character(),
##   City = col_character(),
##   ‘ZIP Code‘ = col_double()
## )

## Warning: 28 parsing failures.
##   row      col               expected     actual                          file
## 16063 ZIP Code no trailing characters 77034-2941 'data/houston_nibrs_2020.csv'
## 21821 ZIP Code no trailing characters 77015-0000 'data/houston_nibrs_2020.csv'
## 29576 ZIP Code no trailing characters 77099-0000 'data/houston_nibrs_2020.csv'
## 56969 ZIP Code no trailing characters 77099-5505 'data/houston_nibrs_2020.csv'
## 56976 ZIP Code no trailing characters 77099-5505 'data/houston_nibrs_2020.csv'
## ..... ........ ..................... .......... .............................
## See problems(...) for more details.

##
## -- Column specification --------------------------------------------------------
## cols(
##   Incident = col_double(),
##   RMSOccurrenceDate = col_character(),
##   RMSOccurrenceHour = col_double(),
##   NIBRSClass = col_character(),
##   NIBRSDescription = col_character(),
##   OffenseCount = col_double(),
##   Beat = col_character(),
##   Premise = col_character(),
##   StreetNo = col_character(),
##   StreetName = col_character(),
##   StreetType = col_character(),
##   Suffix = col_character(),
##   City = col_character(),
##   ZIPCode = col_double()
## )

## Warning: 10 parsing failures.
##   row     col               expected     actual                          file
## 30894 ZIPCode no trailing characters 77071-0000 'data/houston_nibrs_2021.csv'
## 30895 ZIPCode no trailing characters 77071-0000 'data/houston_nibrs_2021.csv'
## 36598 ZIPCode no trailing characters 77043-0000 'data/houston_nibrs_2021.csv'
## 39754 ZIPCode no trailing characters 77045-0000 'data/houston_nibrs_2021.csv'
## 45902 ZIPCode no trailing characters 77048-0000 'data/houston_nibrs_2021.csv'
## ..... ....... ..................... .......... .............................
## See problems(...) for more details.
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   `Case Number` = col_character(),
##   `Occurred From Date` = col_character(),
##   `Occurred Through Date` = col_character(),
##   `Date Reported` = col_character(),
##   Disposition = col_character(),
##   `Disposition Date` = col_character(),
##   `Index Crime Category` = col_character(),
##   `Statute or Ordinance` = col_character(),
##   `Statute or Ordinance Description` = col_character(),
##   `Crime Code` = col_character(),
##   `Crime Code Description` = col_character(),
##   `NCIC Code` = col_character(),
##   `NCIC Code Description` = col_character(),
##   `Domestic Violence` = col_character(),
##   `Crime Location` = col_character(),
##   City = col_character(),
##   Zip = col_character(),
##   `Location Point` = col_character(),
##   `Patrol Division` = col_character()
## )
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   OffenseID = col_double(),
##   FirstDate = col_character(),
##   FirstTime = col_character(),
##   LastDate = col_character(),
##   LastTime = col_character(),
##   ReportDate = col_character(),
##   ReportTime = col_character(),
##   OffenseCode = col_double(),
##   UCRCode = col_character(),
##   IBRCode = col_character(),
##   ATTEMPTvCOMPLETE = col_character()
## )
```

```
## Warning: 3 parsing failures.
##   row         col expected  actual                     file
## 15801 OffenseCode a double S55HRPP 'data/fspd_offenses.csv'
## 35863 OffenseCode a double S55HR   'data/fspd_offenses.csv'
## 39013 OffenseCode a double S55HR   'data/fspd_offenses.csv'
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   OFFENSE_TYPE_ID = col_double(),
##   OFFENSE_CODE = col_character(),
##   OFFENSE_NAME = col_character(),
##   CRIME_AGAINST = col_character(),
```

```
##   CT_FLAG = col_character(),
##   HC_FLAG = col_character(),
##   HC_CODE = col_character(),
##   OFFENSE_CATEGORY_NAME = col_character(),
##   OFFENSE_GROUP = col_character()
## )

##
## -- Column specification --------------------------------------------------------
## cols(
##   `Date Occurred` = col_character(),
##   `Crime Description` = col_character()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   `Date Occurred` = col_character(),
##   `Crime Description` = col_character()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   `Date Occurred` = col_character(),
##   `Crime Description` = col_character()
## )

##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_character(),
##   `Incident Number` = col_double(),
##   `Highest Offense Code` = col_double(),
##   `Occurred Time` = col_double(),
##   `Report Time` = col_double(),
##   `Zip Code` = col_double(),
##   `Council District` = col_double(),
##   PRA = col_double(),
##   `Census Tract` = col_double(),
##   `X-coordinate` = col_double(),
##   `Y-coordinate` = col_double(),
##   Latitude = col_double(),
##   Longitude = col_double()
## )
## i Use `spec()` for the full column specifications.

## Warning: 2 parsing failures.
##    row col               expected actual                      file
## 716672 PRA no trailing characters   526K 'data/Crime_Reports.csv'
## 809259 PRA a double                 CHAR 'data/Crime_Reports.csv'

##
```

```
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_double(),
##   state = col_character(),
##   city = col_character(),
##   NAME = col_character()
## )
## i Use `spec()` for the full column specifications.

## Warning: 792 failed to parse.

## Joining, by = c("Incident", "date", "hour", "class", "description", "offense_count", "Beat", "Premis

## Joining, by = c("Incident", "date", "hour", "class", "description", "offense_count", "Beat", "Premis

## Warning: 37 failed to parse.

## Joining, by = c("date", "time_period")

## Joining, by = c("date", "time_period", "n")

## Joining, by = c("date", "time_period")

## Joining, by = c("date", "time_period", "n")

## Joining, by = c("Date Occurred", "Crime Description")
## Joining, by = c("Date Occurred", "Crime Description")

## Warning: 1 failed to parse.

## Joining, by = c("date", "time_period")

## Joining, by = c("date", "time_period", "n")

## Joining, by = c("date", "time_period")

## Joining, by = c("date", "time_period", "n")

## Joining, by = c("date", "time_period")

## Joining, by = c("date", "time_period", "n")

## Joining, by = c("date", "time_period")

## Joining, by = c("date", "time_period", "n")

## Joining, by = "NAME"

## Joining, by = "NAME"

## Joining, by = c("diff", "jurisdiction")

## Joining, by = c("diff", "jurisdiction")
```

Introduction and Background

Qualified immunity is a court-established doctrine that shields government officials from personal liability for constitutional violations unless the officials violated clearly established laws (https://www.lawfareblog.com/what-qualified-immunity-and-what-does-it-have-do-police-reform). After the killing of George Floyd sparked movements against police violence around the country, many activists directed their attention towards qualified immunity as a subject of reform. Activists argued that qualified immunity prevents police officers from being held accountable for excessive uses of force. Supporters of qualified immunity argued that efforts to limit qualified immunity would prevent police officers from effectively carrying out their job out of fear of frivolous lawsuits. In this statistical report, we aim to provide preliminary data-driven insights on the effects of recently passed qualfied immunity reform on violent and property crime rates in major urban jurisdictions.

Four states have passed measures to limit qualified immunity: Colorado, Connecticut, New Mexico, and New York. Of those states, Colorado passed its reform the earliest, with its measure taking effect June 19 of 2020. As a result, we decided to analyze Colorado crime data to determine the effects of qualified immunity on crime rates. In particular, our research question was as follows: Was the passage of qualified immmunity reform in Colorado in June of 2020 correlated with significantly larger proportional increases in average daily violent and property crime incidents than increases in control jurisdictions? Because statewide incident-level data for 2020 and 2021 YTD from Colorado was not publicly available, we further narrowed the scope of our analysis to Denver and Colorado Springs, the two largest jurisdictions within Colorado.

Although we attempt to esttablish some level of causation in this study through the use of something similar to a synthetic control method, we lack the volume of observational data needed to successfully establish causation. In particular, we are missing observations on several key lurking variables, including the effects of COVID on poverty rates in each jurisdictions, 2020 and 2021 census data, community attitudes towards policing as a result of the George Floyd protests, amidst several others. Much of this data will only be released a few years from now, severely limiting the contours of this analysis. However, due to the prescience of the qualified immunity question and the need for data within the debate, we decided to produce this preliminary report to at least illustrate the plausible effects of qualified immunity on crime rates in Colorado. None of the findings in this report should be interpreted as demonstrating a conclusive causal relationship between qualified immunity reform and crime.

Methodology

In this study, we constructed a synthetic control model as described by Alberto Abadie in his article "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." In a synthetic control model, researchers create a weighted average of jurisdictions and data points with the goal of minimizing the distance between the weighted average and the true jurisdiction's pre-treatment predictor and response values. The synthetic control model has the advantage of systematically identifying the strongest control jurisdictions based on predictor numbers over time, as opposed to simply manually identifying a geographically close jurisdiction and claiming the existence of sufficient connection to isolate the effects of the policy intervention, thus removing the effects of researcher bias from the analysis. Generally, the synthetic control method is then followed by a series of placebo synthetic control constructions to determine if the post-treatment to pre-treatment RMSPE (root mean squared prediction error) ratio for the treated jurisdiction is extreme compared to placebo synthetic controls (https://mixtape.scunning.com/synthetic-control.html). If employed successfully, the synthetic control method can successfully indicate the existence of a causal connection between different variables.

In practice, however, we faced critical data limitations that made it impossible to run the most statistically robust version of the synthetic control method (which is also why we refuse to arrive at any causal conclusion). Because we lacked widely available 2020-21 post-treatment crime and census data (we only had that data for very specific jurisdictions), we could not create a series of placebo synthetic controls with post-treatment RMSPE values to compare the RMSPE ratio for the treated jurisdiction with. In addition, we only had access to a very small number of time periods (2011-19) and included an enormous sample of jurisdictions within the donor pool (many of which were substantially different from the treated jurisdictions) due to the rarity by which we could obtain data from police departments, increasing the bias bound by a large amount according to Abadie's article. To account for some of these severe limitations, we primarily utilized the synthetic control methodology to identify similar jurisdictions to Denver or Colorado Springs and to weigh four to five of those jurisdictions to create a control for comparative purposes.

Because no database of city predictors and violent/property crime rates over 2011-2019 previously existed, we created a new database from ACS census data from 2011 to 2019 recording each jurisdiction's name, single female-led family household percentage, percentage of people who were high school graduates or higher, percentage of people who lived in the same house they lived in a year ago, percentage of people who are over 18, percentage of people who are white, percentage of people who are self-employed, unemployment rate, median income, child poverty rate, and the percentage of housing units occupied by their owners. Each of these predictors had been identified as possible or significant predictors of crime within metropolitan and nonmetropolitan counties in a different study published by researchers Wells and Weishelt (Explaining Crime in Metropolitan and Non-Metropolitan Counties). We then combined the city predictor data with violent crime, property crime, and population statistics from UCR Crime in the United States fact tables. Our final full database included over 83000 observations and 20 variables, each observation representing a jurisdiction at a particular year. We then filtered the database to only include cities above 50000 in population to remove small rural jurisdictions that would not likely match the dynamics of more urban areas like Denver. Additionally, we removed jurisdictions with missing data on violent/property crime rates or missing yearly data.

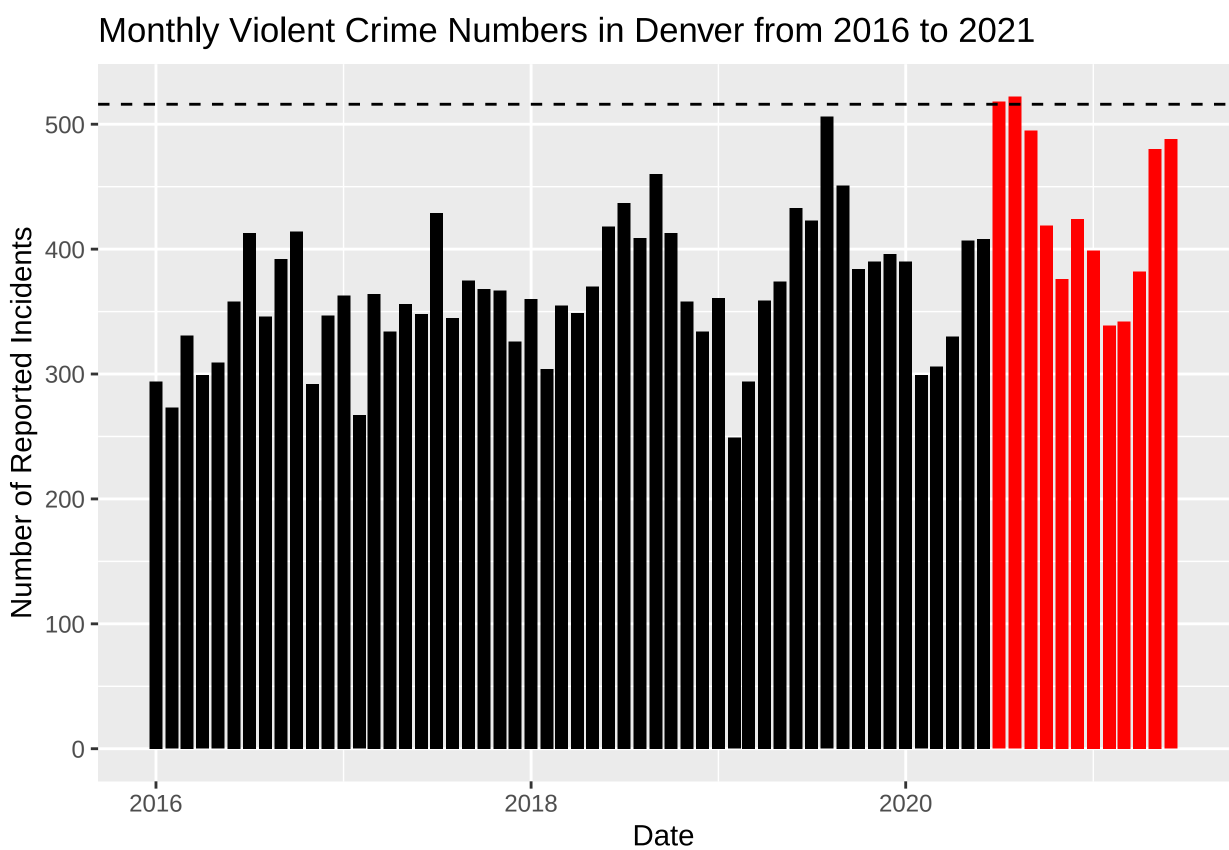
Since publicly available data was not available for 2020-2021 from either the UCR or the ACS, we utilized the Synth package to create a synthetic control model for Denver and Colorado Springs from 2011-2019. We optimized the synthetic control model for 2016 to 2019 to obtain jurisdictions that could follow the more recent trends in both Denver and Colorado Springs. We then identified roughly the top 5 jurisdictions with the highest weights and reran the synthetic control model with only those jurisdictions to recalculate the weights. With those identified control jurisdictions, we submitted requests for incident-level crime data from those departments for 2019-2021. When those requests were either unanswered or denied (as in the case of Ann Arbor Police Department), we removed the city from the synthetic control model and reran the model until we obtained at least police departments with accessible incident level crime data whose plot looked at least somewhat similar to the plots of the treated jurisdictions (Denver and Colorado Springs).

We subdivided all incident-level crime data into property and violent crimes based on UCR definitions. In particular, murder and nonnegligent homicide, aggravated assault, robbery, and forcible rape (including sexual assault with an object, fondling, and forcible sodomy) were identified as violent crimes. We categorized larceny charges, burglary, property damage/destruction, arson, shoplifting, pocket-picking, and motor vehicle theft charges as property crimes. We then calculated the daily numbers of violent and property offenses for June of 2019 to June of 2020 (before qualified immunity reform) and June of 2020 to June of 2021 (after qualified immunity reform) in control and treated jurisdictions. We then subtracted the daily numbers of violent and property offenses in the 2019-20 time period from the 2020-21 time period and divided by the total number of violent or property offenses in the 2019-20 time period to make the daily numbers of violent and propery crimes proportionate to each jurisdiction's respective crime numbers. Finally, we created a bootstrapped null distribution assuming no true difference between the increases of the synthetic jurisdiction and Denver or Colorado Springs and calculated a p-value based on the probability of observing the real difference or greater between Denver/Colorado Springs and the respective synthetic control difference.

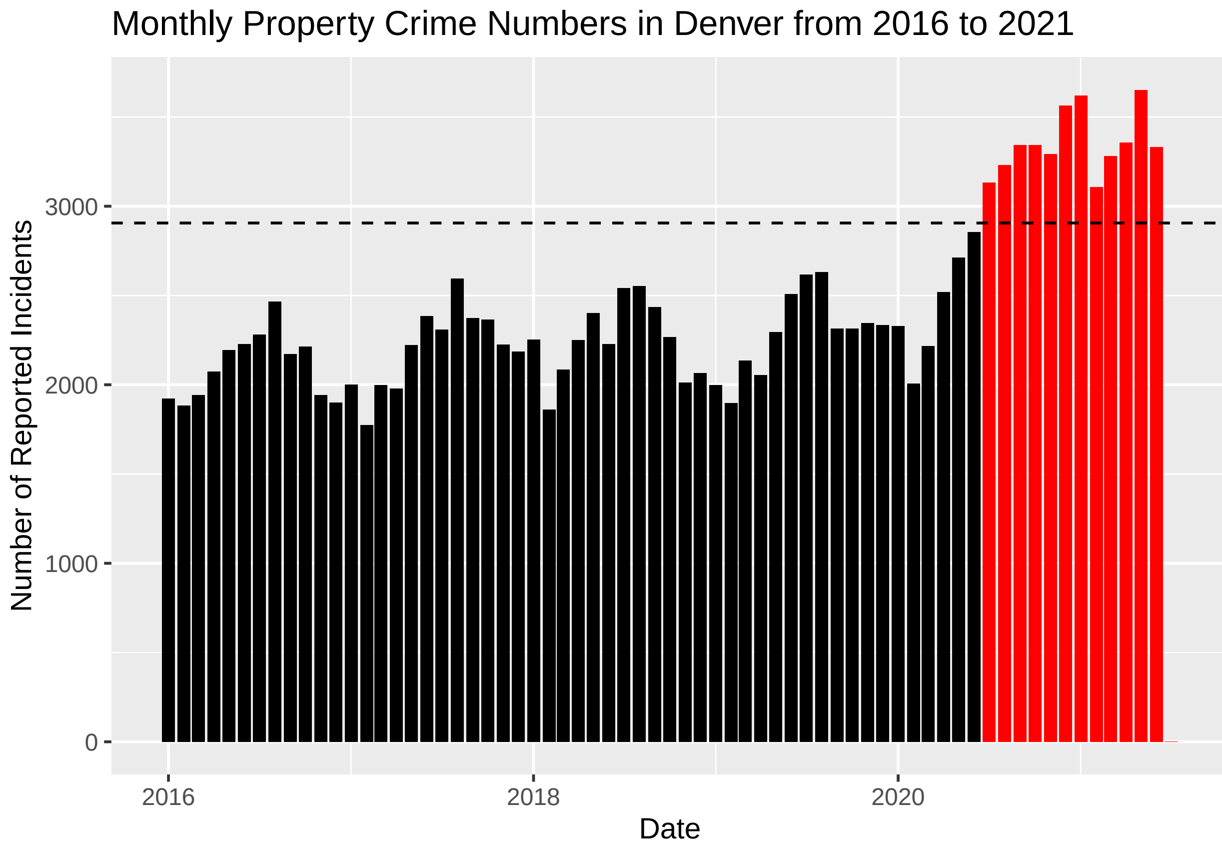
This methodology had a few other critical limitations. First, because of the lack of 2020 and 2021 census data, we could only identify jurisdictions similar to either Denver or Colorado Springs until 2019 which is no guarantee that those similarities in control predictors continued until 2020 and 2021. In addition, since we did not have access to UCR data for 2020 and 2021, we had to use 2011-2019 weights in 2020 and 2021 calculations, which may extrapolate beyond the capabilities of the synthetic control model, since the trends between the predictor variables and the responsive crime rates may not continue into 2020 and 2021 (which also introduced other variables, such as the George Floyd protests, that influenced crime rates). Second, because of the several denied requests, we had to employ a form of convenience sampling in order to successfully carry out the study, as many jurisdictions refused to allow us to download their incident-level data, as in the case of Ann Arbor Police Department and Clarksville Police Department. Despite this, the plots comparing the synthetic control property/violent crime rates with the Denver and Colorado Springs property/violent crime rates look strong enough to indicate that the synthetic control methods at least partially match the crime trends of the treated jurisdictions even when certain jurisdictions were removed due to lack of data. Third, because we needed to determine if the increases between the 2019-20 time period and 2020-21 time period were significantly greater than increases in control jurisdictions, we were forced to employ a test where we subtracted daily crimes in one time period from daily crimes in another time period. This may have exaggerated the standard deviation of violent and property crimes, since daily fluctuations in crime do not remain constant over the course of a year. The test may have been more successful on a monthly level, but we could not employ such a test, since the number of observations would be too low.

Exploratory Data Analysis

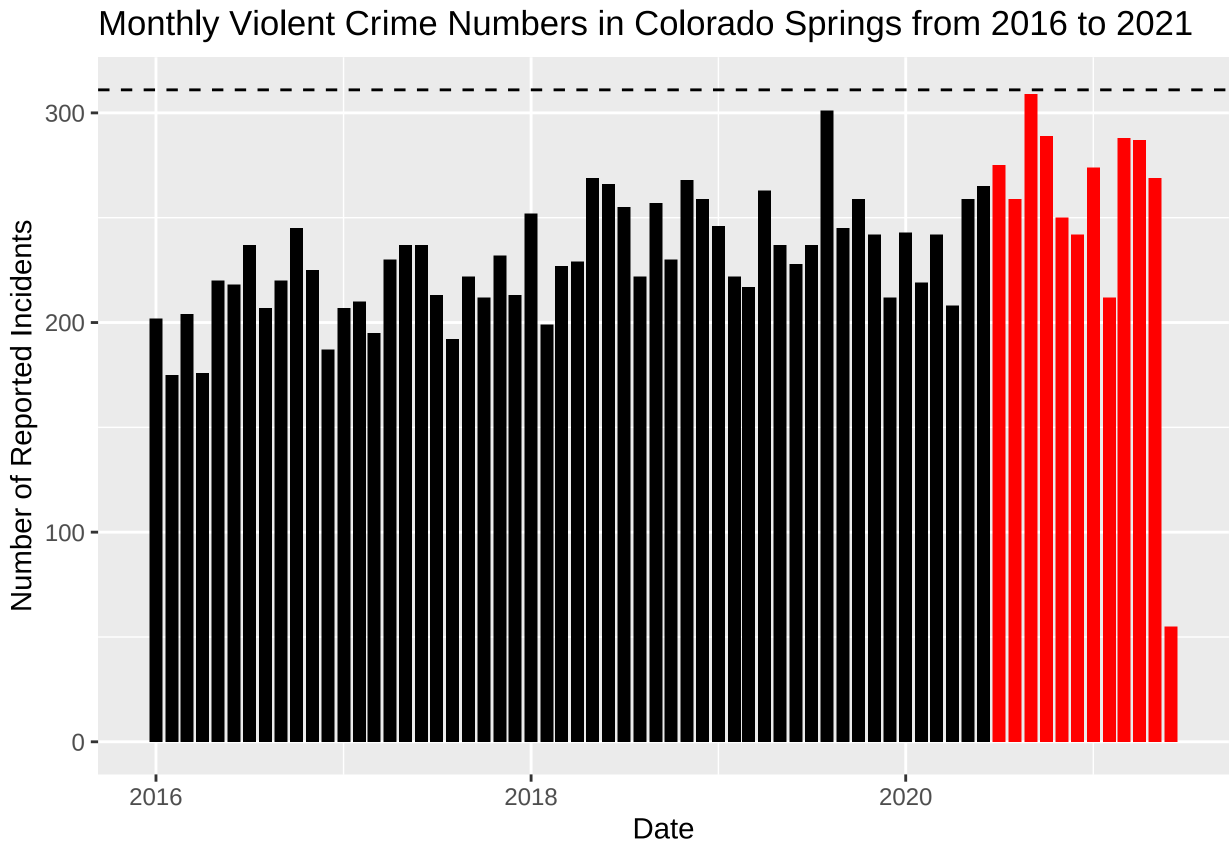
We began by visualizing daily violent crime and property crime numbers in both Denver and Colorado Springs.



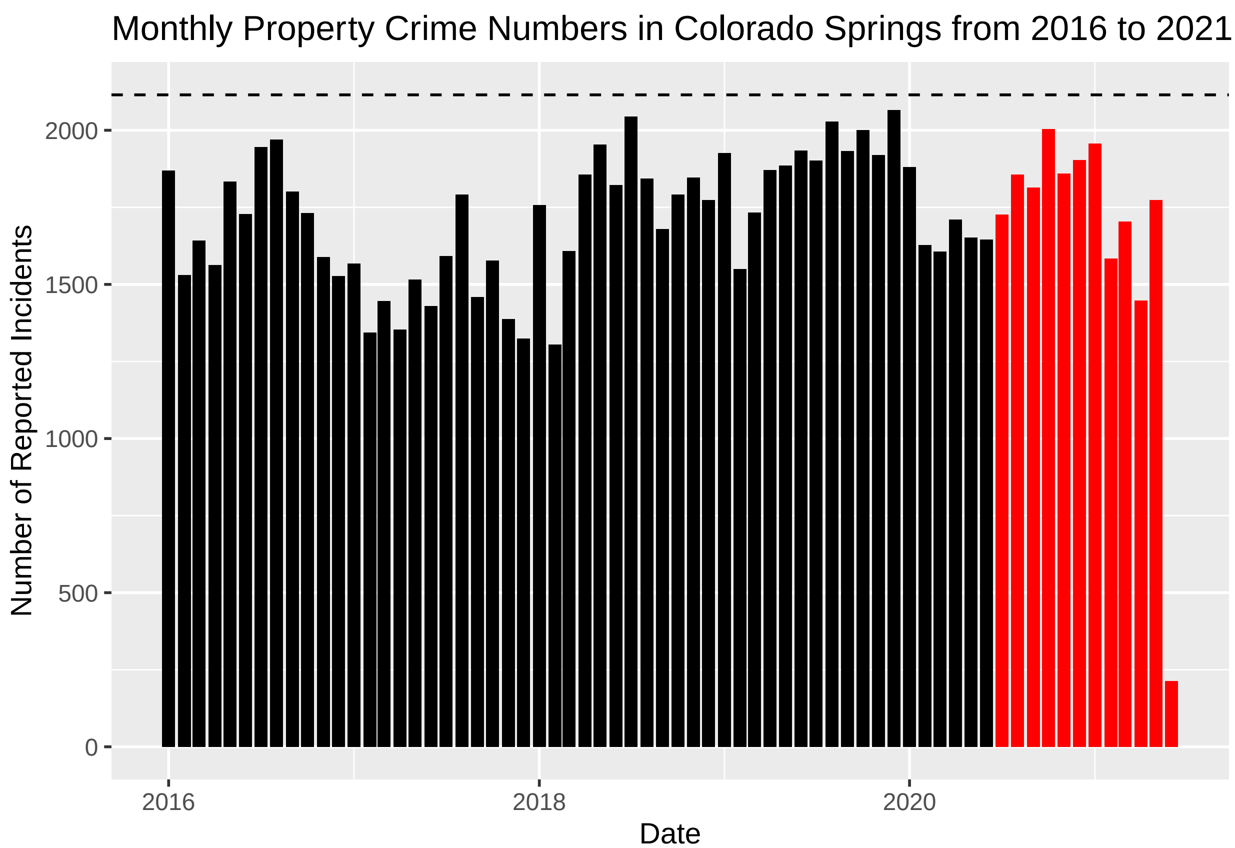
The black dotted line represents 10 reported incidents above the maximum number of violent crime offenses in a single month in the four years prior to passage of qualified immunity reform. Although 2020 did have 2 months surpass the maximum number of violent crime offenses in a single month, there does not appear to be a significantly greater number of violent crimes after the passage of qualified immunity reform compared to qualified immunity reform beforehand. The graph also indicates the importance of controlling the effects of seasonal shifts on crime; since violent crimes tend to increase during the summer, we cannot employ a methodology of comparing the first half of 2020 to the second half of 2020 because the first half of 2020 encompassed less of the high-crime summer months.



On the other hand, for Denver property crimes, almost every month after the passage of qualified immunity reform far surpassed the previous maximum number of property offenses in a single month. The bar graph provides preliminary evidence that Denver property crimes increased substantially after the passage of the police accountability bill in June of 2020, although the bar graph also appears to depict a rise in property crime that began before the passage of the bill (the spike beginning in roughly April of 2020 which already matched the highs of the summer months beforehand). Unfortunately, it is difficult to disentangle the effects of normal seasonal shifts and the COVID recession from the effects of the police accountability bill (if there were any).



Similarly to Denver, Colorado Springs depicted no readily apparent increase in violent crime numbers after the passage of qualified immunity (depicted in red bars).

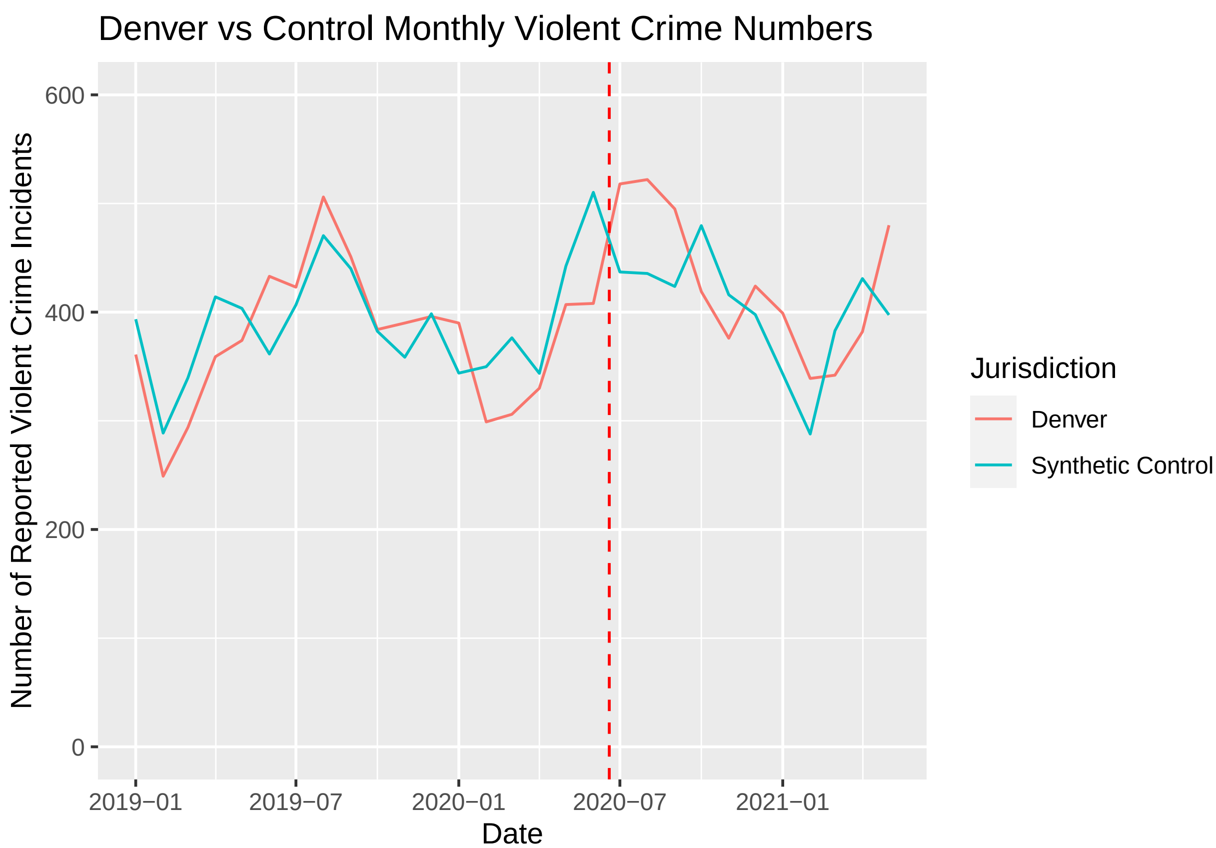


Diverging from Denver trends, however, Colorado Springs lacked the large rise in property crime numbers that Denver saw after the passage of qualified immunity. The graphs provide some preliminary evidence that the rise in property crime in Denver may have been due to other factors unrelated to the passage of qualified immunity (such as municipal budget issues or worse COVID effects there).

```
## Joining, by = c("date", "jurisdiction")
## Joining, by = c("date", "jurisdiction")
```

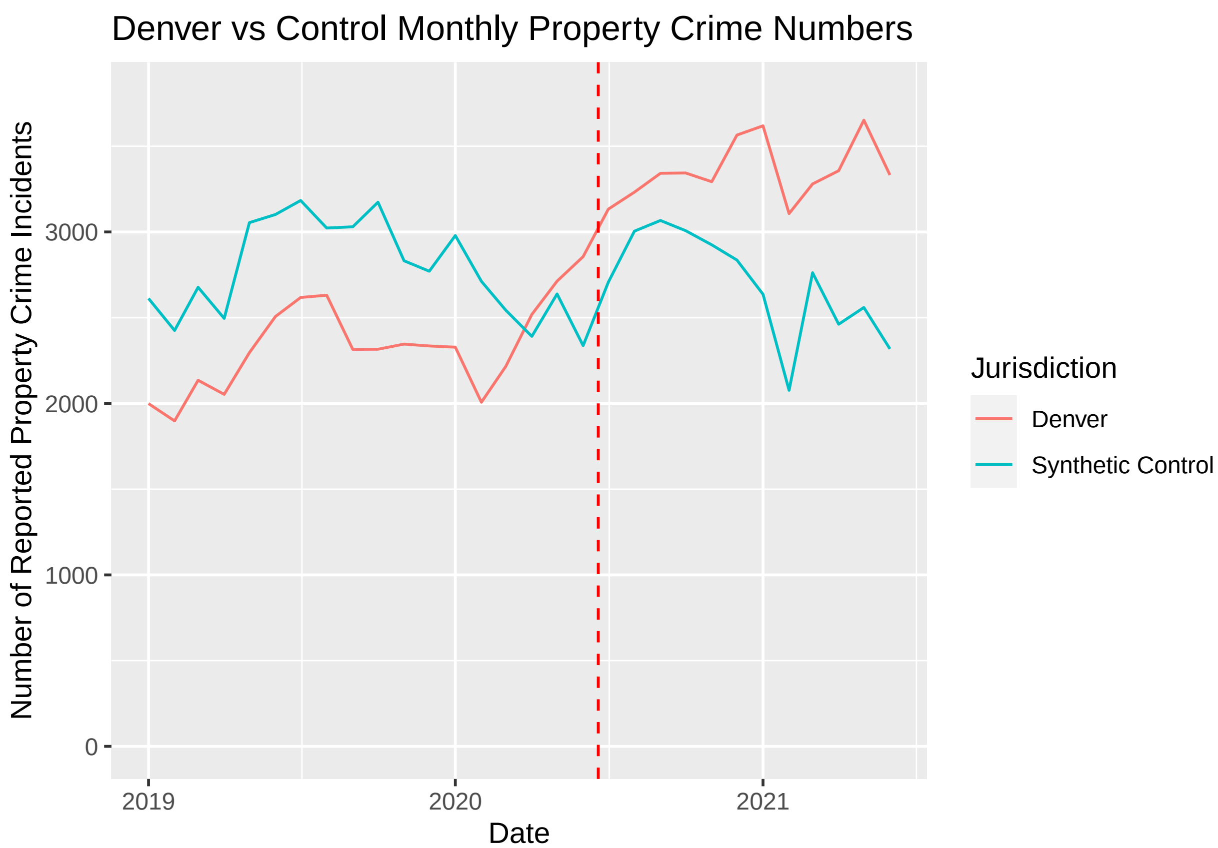
We then directly compared Denver's violent and property crime monthly numbers with synthetic control numbers from 2019 to 2021 (the only publicly available data).

```
## Joining, by = c("my", "n", "jurisdiction")
```



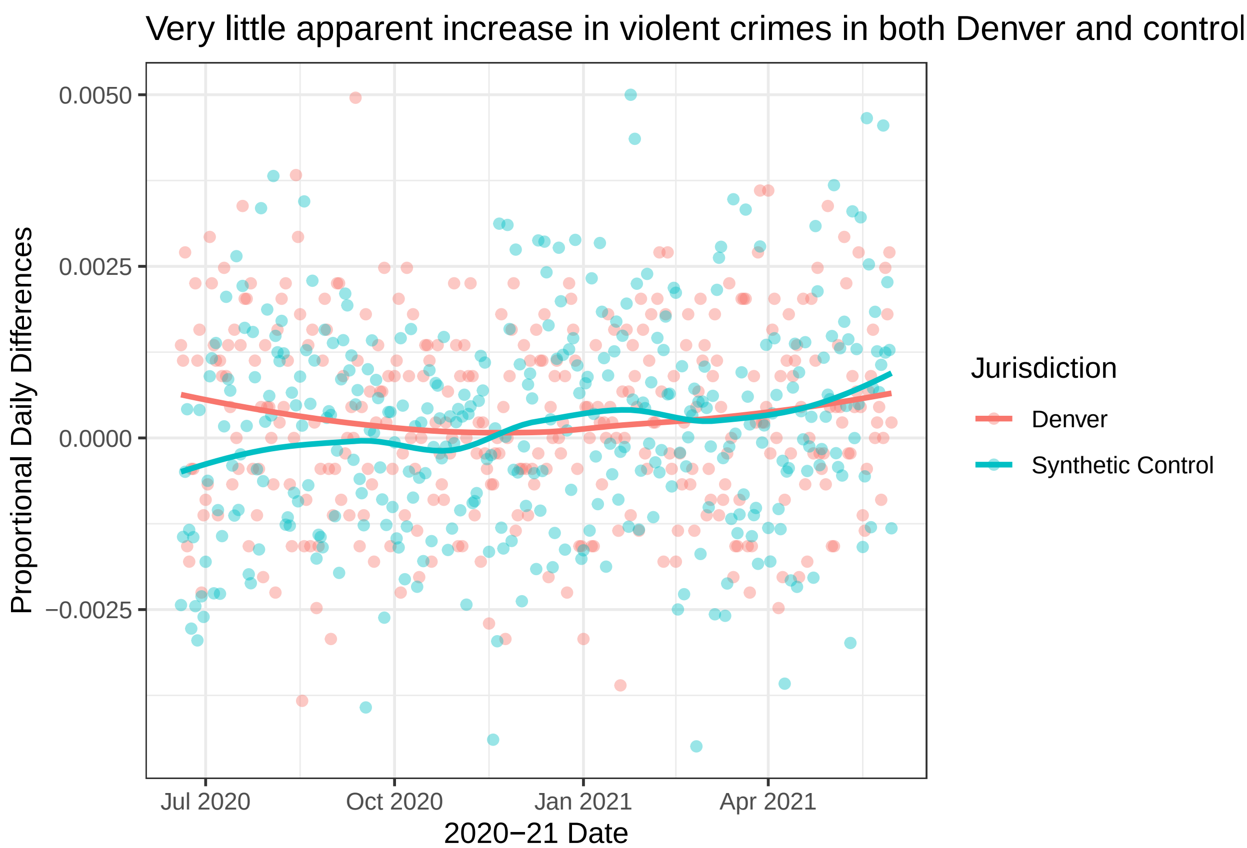
When compared to synthetic control monthly numbers, the graph appears to suggest that Denver's changes in number of monthly violent crime incidents did not vary significantly from similar jurisdictions in terms of violent crime characteristics.

```
## Joining, by = c("my", "n", "jurisdiction")
```

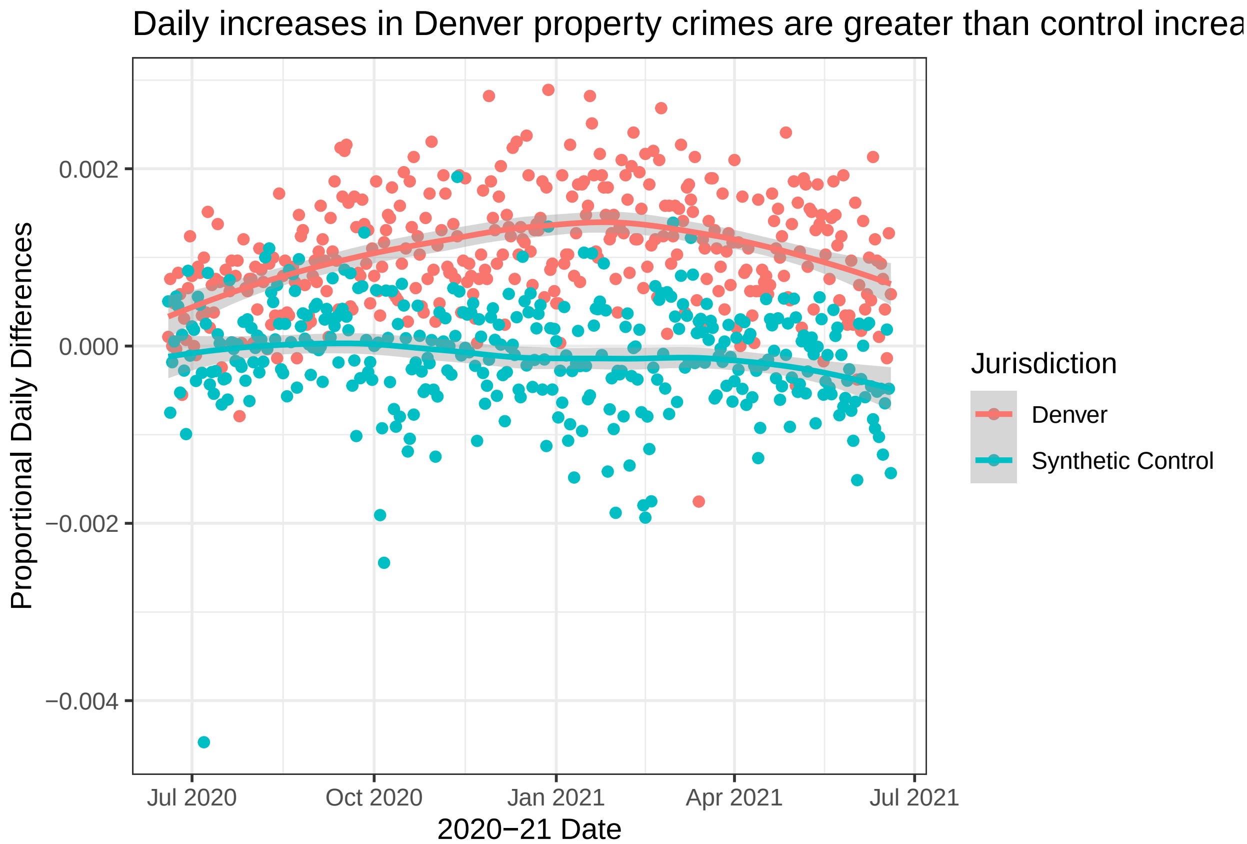


On the other hand, Denver's increase in property crime numbers was significantly greater than increases in other jurisdictions. However, the graph comparing property crime numbers between Denver and the synthetic control unit is also much weaker than the violent crime unit, with the two lines significantly deviating from each other even before the passage of qualified immunity reform.

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



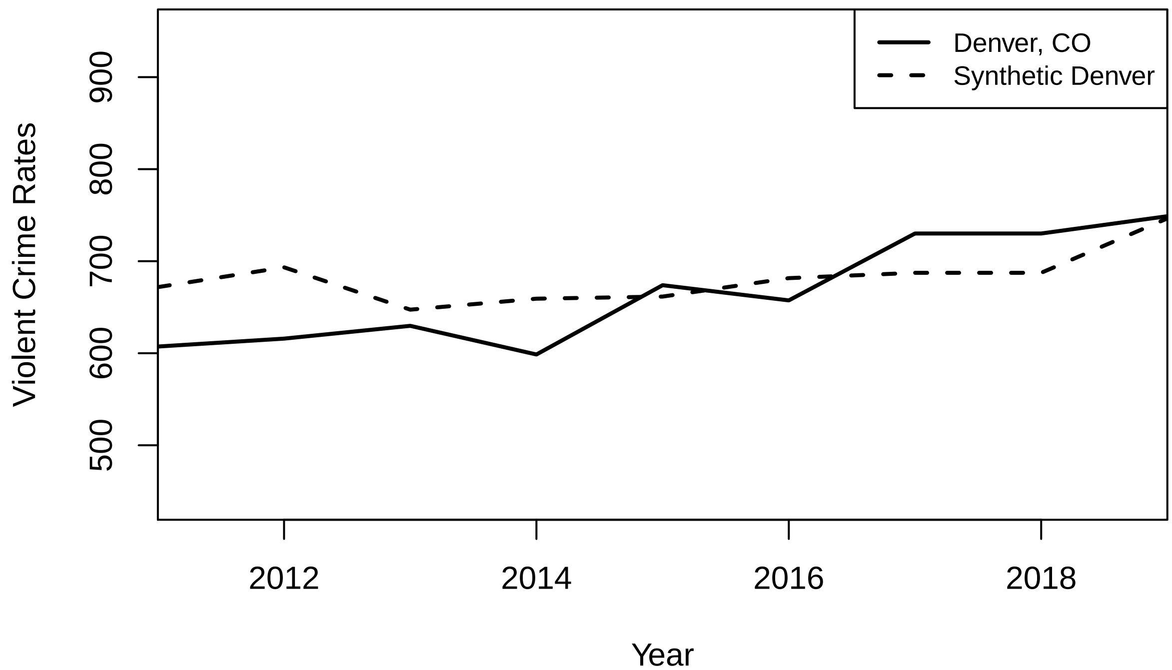

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



Analysis

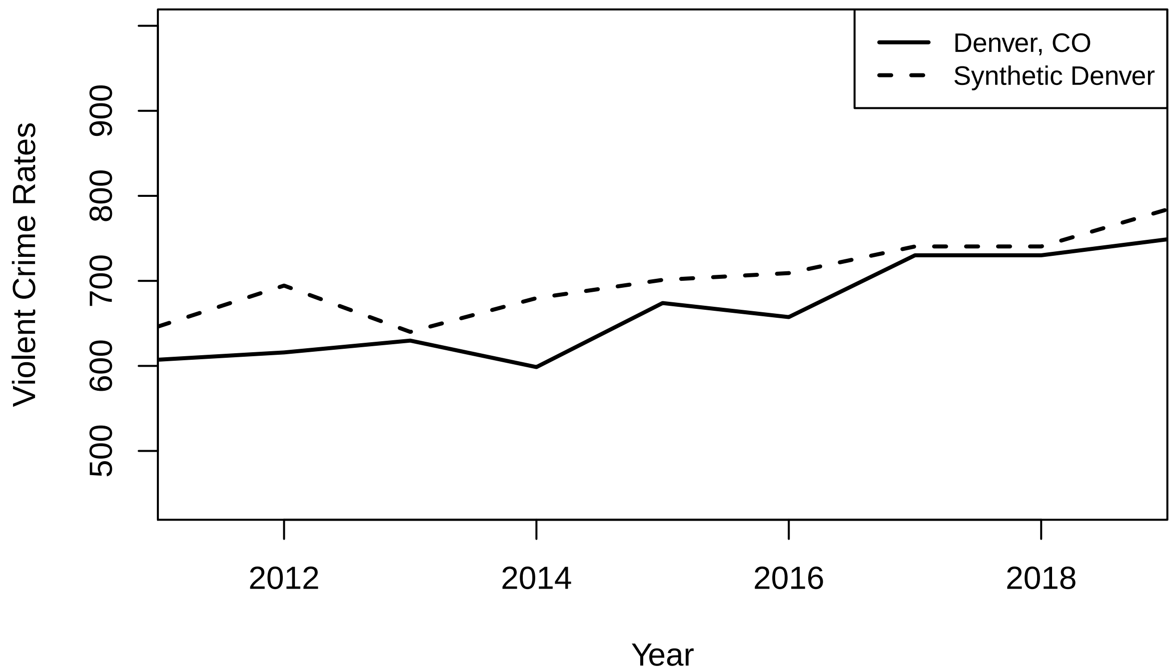
Synthetic Control Modelling

We began by creating the database of all census cities/towns within America from 2011-2019 that had both UCR and ACS data. From there, we filtered the database to only include cities greater than 50000 in population to identify roughly 500 similar jurisdictions to Colorado Springs and Denver. Using each jurisdiction's violent and property crime rates, we created synthetic jurisdictions based on the weighted averages of all 500 jurisdictions (weights determined through proximity to predictor and response values of the treated jurisdictions). The resulting graphs comparing the true treated units with their synthetic control models pre-treatment over 2011-19 are shown below.

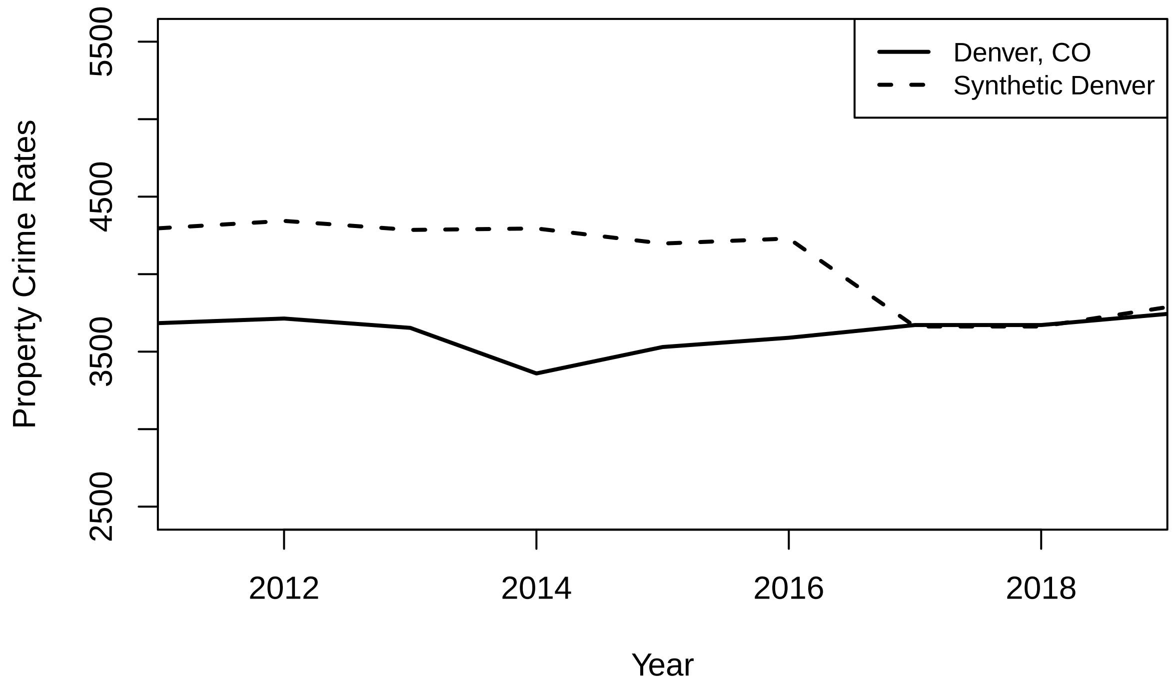


The first graph comparing the Denver annual violent crime rates with the synthetic control model's is relatively strong, with the synthetic control model's violent crime rates roughly matching some of the trends and the level of Denver's violent crime rate pre-treatment. There are some imperfections; for instance, Denver's random yearly fluctuation isn't fully matched each year, although Denver's overall trend is roughly captured in the model.

Due to limitations in obtaining 2020 and 2021 data, I limited the calculation of the weighted average to the top 4 jurisdictions and recalculated the synthetic model to obtain new weights (the 5th jurisdiction, Ann Arbor, denied our FOIA request). The true fit of the model to pre-treatment Denver is depicted below:

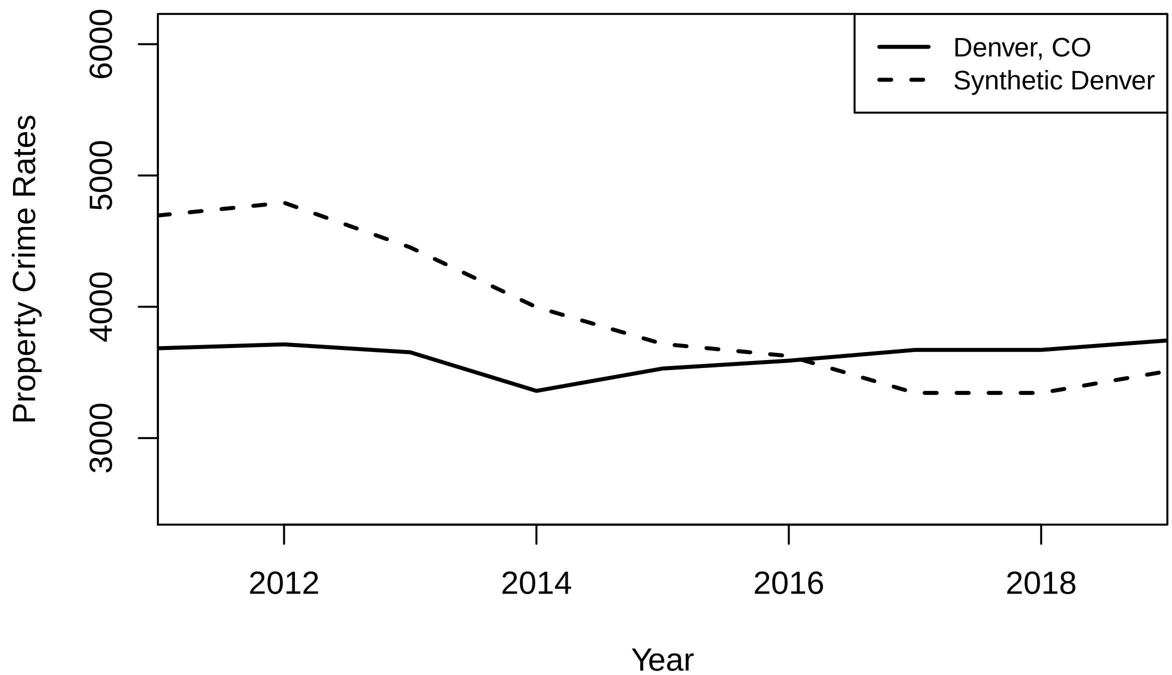


The synthetic of the 4 cities is slightly higher in terms of level but still follows the trend of true pre-treatment Denver roughly correctly. Thus, we found the synthetic control to be similar enough to the real Denver to safely proceed with statistical testing.



On the other hand, the plot comparing pre-treatment Denver and the synthetic control model for Denver do not appear to be on the same level prior to 2017. From 2017 to 2019, the plot is roughly correct in comparing property crime rates, but the significant drop in the synthetic model from 2016 to 2017 indicates that the model may be somewhat limited in its explanatory potential and ability to cast a counterfactual for Denver.

Excluding Ann Arbor, we refitted the synthetic control model with the 5 most highly weighted jurisdictions.



The synthetic of the top 5 cities of the previous synthetic control model is still significantly off in terms of following the trends of Denver in terms of property crimes. This may reflect crucial limitations in the ability of the synthetic control model to account for Denver property crime trends.

Statistical Bootstrapping Simulations

After collecting data from each of the 4-5 jurisdictions identified in each test as well as Denver and Colorado Springs, we calculated the daily differences in violent and property crime between the June 19, 2019 to June 18, 2020 time period compared to the June 19, 2020 to June 19, 2021 time period (the first time period also had an extra day from the leap year). In particular, we corresponded the dates so that the number of violent crimes on June 19, 2019 was subtracted from the number of violent crimes on June 19, 2020 and created a dataset of these differences in violent and property crime numbers. These differences were then divided by the total number of violent or property crimes in the first period of time. We divided by the total number of crimes in the previous period as opposed to the population in order to account for jurisdictions which began from already-high crime rates and the proportionately smaller increase in crime rate that the same absolute increase in crime woudl entail.

To calculate the synthetic control differences for comparison, we used the weights in the previous section and multiplied them by the proportional daily differences in crime between the two periods. We then summed up the the proportional daily differences and joined the two datasets together. We then used bootstrapping to create a null distribution of 10000 differences in mean centered at 0 and determined if the probability of observing the difference between the mean proportional average daily increase in Denver or Colorado Springs with the mean proportional average daily increase in the synthetic control or greater was low enough to justify concluding that Denver or Colorado Springs' increase in violent crime was significantly greater than control jurisdictions.

Denver Violent Crime Tests A table of the synthetic control jurisdictions for Denver violent crimes with weights is shown below:

NAME	weight
Seattle, Washington	0.496466442
Fort Smith, Arkansas	0.487715311
Champaign, Illinois	0.011537542
Houston, Texas	0.004280705

Because the synthetic control jurisdiction gives weight to Houston, we only utilized June 2020 to May 2021 data and June 2019 to May 2020 data for this test. Houston's publicly available data unfortunately concludes after May 31, 2021, meaning we cannot utilize the extra 20 days in June for this test.

$H_0 : \mu_{Denver} = \mu_{Synthetic}$. The true mean daily proportional difference in number of violent offenses between the June 2020 to May 2021 time period compared to the June 2019 to May 2020 time period in Denver, CO is equal to the true mean daily proportional difference in number of violent offenses between the two time periods in the synthetic control model.

$H_A : \mu_{Denver} > \mu_{Synthetic}$. The true mean daily proportional difference in number of violent offenses between the June 2020 to May 2021 time period compared to the June 2019 to May 2020 time period in Denver, CO is greater than the true mean daily proportional difference in number of violent offenses between the two time periods in the synthetic control model.

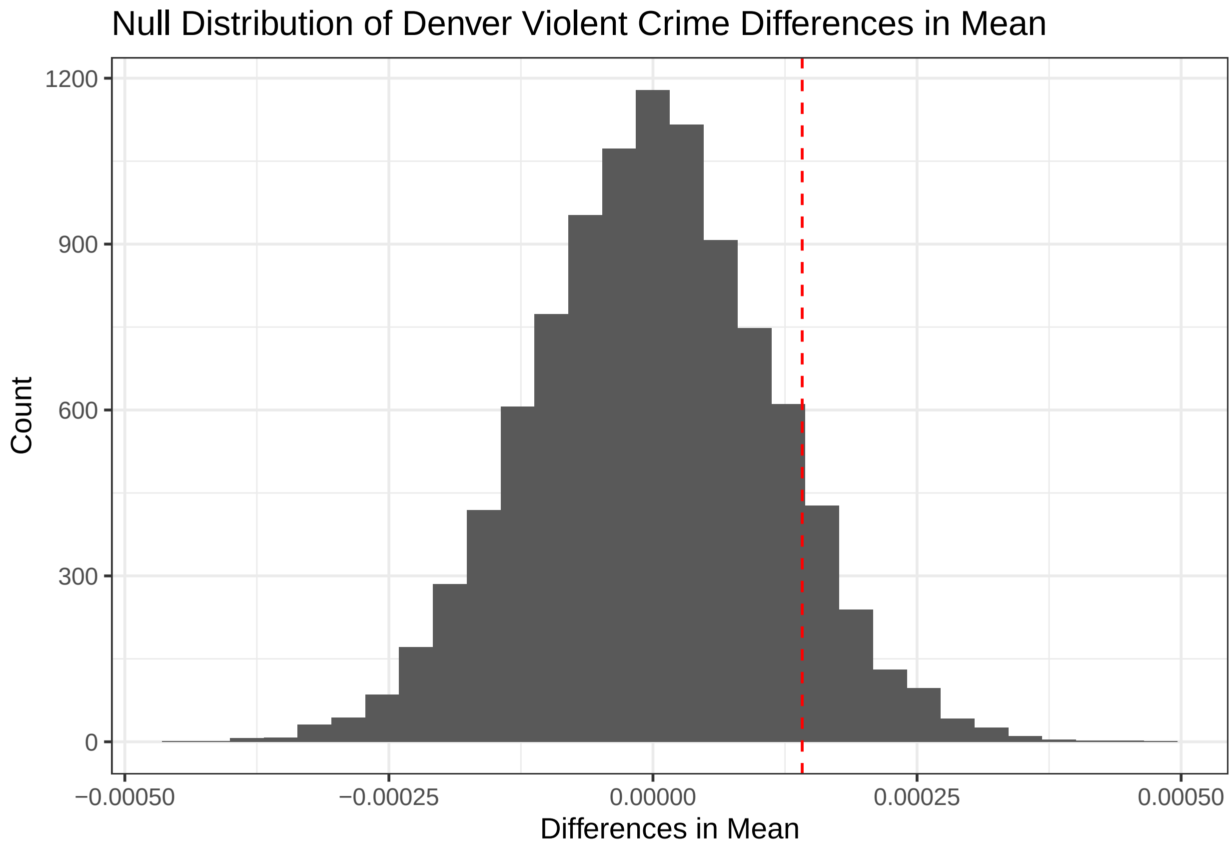
$\alpha = 0.01$

Because t-test conditions may not be fully met, we utilized a bootstrapped simulation.

```
## [1] 0.1025
```

The null distribution (with the red line displaying where the observed difference in mean between Denver and the synthetic control model lies) is graphically shown below:

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



Because the p-value of 0.1025 is greater than a reasonable alpha level of 0.01, we fail to reject the null hypothesis. The data does not provide sufficient evidence at the 1%, 5%, or 10% level that Denver's average daily increase in violent crimes from the 2019-20 time period to the 2020-21 time period is significantly greater than the synthetic control's average daily increase in violent crimes.

We also conducted a monthly difference in difference test using the synthetic control model as the "control" jurisdiction, since the graph modeling the trends of the synthetic control graph with true Denver trends indicated the possibility of parallel yearly violent crime trends between the synthetic control model and Denver, although the levels of the two models did not exactly match. By utilizing monthly data and linear modeling, we decreased the influence of daily crime fluctuations on the results while simultaneously retaining sufficient data points to draw some statistical conclusions.

```
## Warning in Ops.factor(treated, time): '*' not meaningful for factors
```

```
## # A tibble: 4 x 5
##   term           estimate std.error statistic  p.value
##   <chr>             <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)       390.       13.9    28.1   6.27e-34
## 2 time1              12.7      22.5     0.563 5.76e- 1
## 3 treated1          -14.7      19.6    -0.748 4.58e- 1
## 4 time1:treated1     38.7      31.9     1.21  2.30e- 1
```

The interaction variable for time and treated reflects the difference in difference estimate. Because the p-value of 0.230 far exceeds a reasonable alpha level of 0.01, we fail to reject the hypothesis that the passage of qualified immunity in June 19, 2020 in Denver did not correspond to an increase in violent crime that outpaced other control jurisdictions.

Denver Property Crime Tests The following jurisdictions and weights were utilized to construct the synthetic control for Denver property crimes:

NAME	weight
Austin, Texas	6.935988e-01
Champaign, Illinois	2.188792e-01
Seattle, Washington	8.747524e-02
Fort Smith, Arkansas	4.654831e-05
Houston, Texas	1.805219e-07

Since Houston had such a low weight and lacked significant data (such as June statistics), I decided to exclude Houston from the analysis and run the bootstrapping with data from June 19, 2020 to June 19, 2021, accounting for a whole year of qualifeid immunity reform.

$H_0 : \mu_{Denver} = \mu_{Synthetic}$. The true mean daily proportional difference in number of property offenses between the June 2020 to June 2021 time period compared to the June 2019 to June 2020 time period in Denver, CO is equal to the true mean daily proportional difference in number of violent offenses between the two time periods in the synthetic control model.

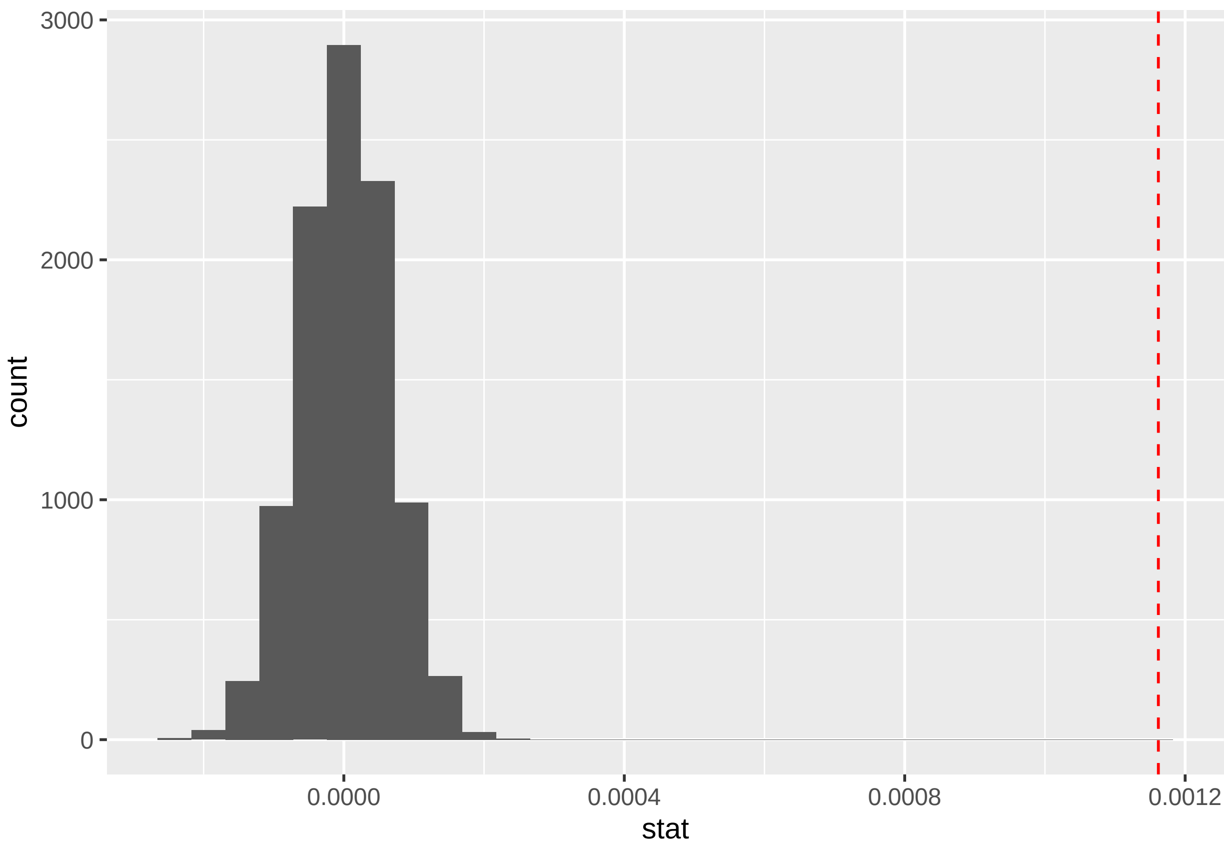
$H_A : \mu_{Denver} > \mu_{Synthetic}$. The true mean daily proportional difference in number of violent offenses between the June 2020 to June 2021 time period compared to the June 2019 to June 2020 time period in Denver, CO is greater than the true mean daily proportional difference in number of violent offenses between the two time periods in the synthetic control model.

$\alpha = 0.01$

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

A histogram of the null distribution is shown below. The red line indicates that the true difference in mean between Denver and the synthetic control's increases in property crime is well above the hypothesized null distribution centered at 0.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Because the p-value of 0 is far less than a reasonable alpha level of 0.01, we reject the null hypothesis. The data provides sufficient evidence to indicate that Denver's average daily proportional increase in property crime after the passage of qualified immunity was significantly greater than control cities' increase in property crime over the same time period.

We did not employ a monthly difference in difference test because the parallel trends assumption is clearly violated, since the graph comparing the synthetic control trends with the true Denver trends is not parallel, especially from 2016-2017.