Statistical Report on Qualified Immunity

Setup

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 within civil court. 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 is as follows: Was the passage of qualified immunity 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 establish some level of causation in this study through the use of 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 jurisdiction, 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 the present 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 utilized a synthetic control method as described by Abadie in his article "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." In a synthetic control method, 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 method has the advantage of systematically identifying the strongest control jurisdictions based on predictor numbers over time, as opposed to 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 those of 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. 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, significantly increasing the potential for bias. To continue the data analysis under these limitations, we primarily utilized the synthetic control methodology to identify similar jurisdictions to Denver or Colorado Springs and to provide weights for some of those jurisdictions.

Because no database of city predictors and violent/property crime rates over 2011-2019 previously existed, we created a new database from scratch. Using ACS census data from 2011 to 2019, we recorded each jurisdiction's name, single female-led family household percentage, percentage of the population who graduated high school, percentage of the population who lived in the same house they lived in a year ago, percentage of the population who were over 18, percentage of the population who were white, percentage of the population who were 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 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 likely not 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 most 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 four 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 2019-2021 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 of property, arson, shoplifting, pocket-picking, and motor vehicle theft charges as property crimes. We 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 daily increases of the synthetic jurisdiction compared to 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 based on the null distribution.

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, 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 employed a form of convenience sampling in order to successfully carry out the study, as many jurisdictions refused to allow us to obtain 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 controls 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 method may have substantially 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 did not have enough monthly difference data to successfully arrive at statistical conclusions through simulation.

Exploratory Data Analysis

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

Monthly Violent Crime Numbers in Denver from 2016 to 2021
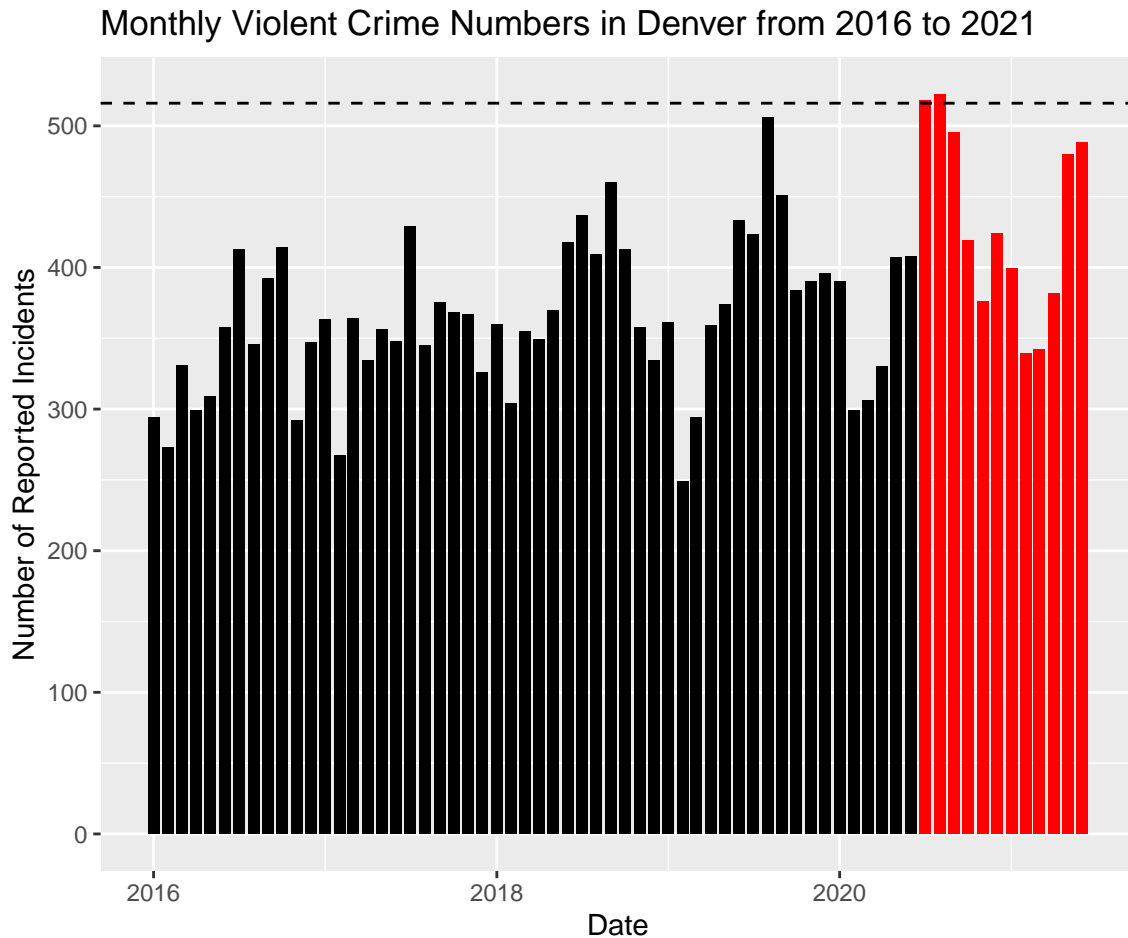

The black dotted line represents 10 reported incidents above the maximum number of violent crime offenses in a single month in the four years prior to passage of qualified immunity reform. Based on the graph, 2020 did appear to have significant increases in violent crime (with 2 months surpassing the previous records for the highest number of violent crimes in a single month). However, when observing the previous 5 years of crime, the crime numbers after the passage of the police accountability bill may be interpreted as part of a steady yearly increase in violent crime numbers from 2016 to 2021. The graph also indicates the importance of controlling the effects of seasonal shifts on crime; violent crime numbers tend to peak during the hot summer months and decrease during the winter.

## Monthly Property Crime Numbers in Denver from 2016 to 2021
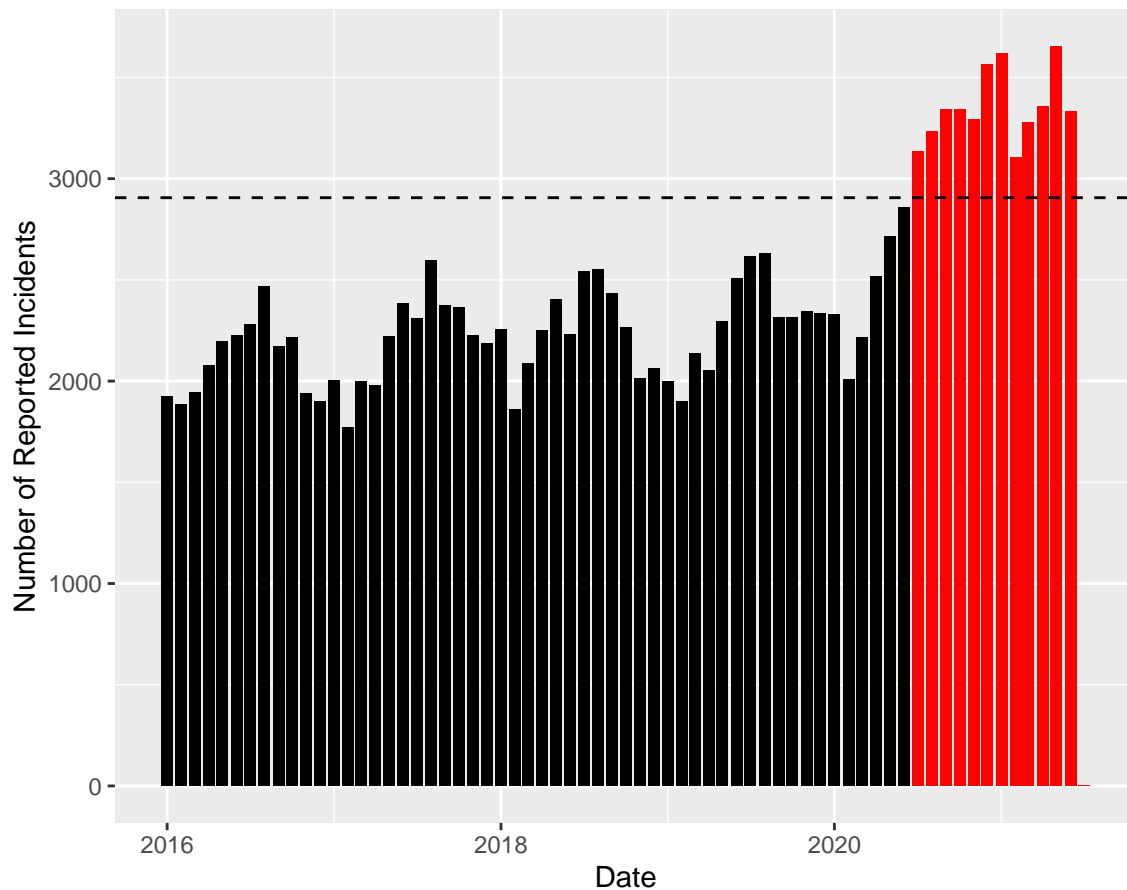
The trend we observed in Denver violent crimes is remarkably turned on its head when examining property crimes. Almost every month after the passage of qualified immunity reform, we observed property crime numbers that far surpassed the previous maximum number of property offenses in a single month. 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 (April of 2020 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 possible effects of the police accountability bill.

If the qualified immunity measures in the police accountability bill *caused* increases in property crime in Colorado, we would expect to see major property crime increases in jurisdictions throughout Colorado. To determine if Denver trends were replicated elsewhere, we included the following two graphs from Colorado Springs, despite the fact that our preliminary analysis has not gotten far enough to successfully compare Colorado Springs with a synthetic control.

Monthly Violent Crime Numbers in Colorado Springs 2016–2021
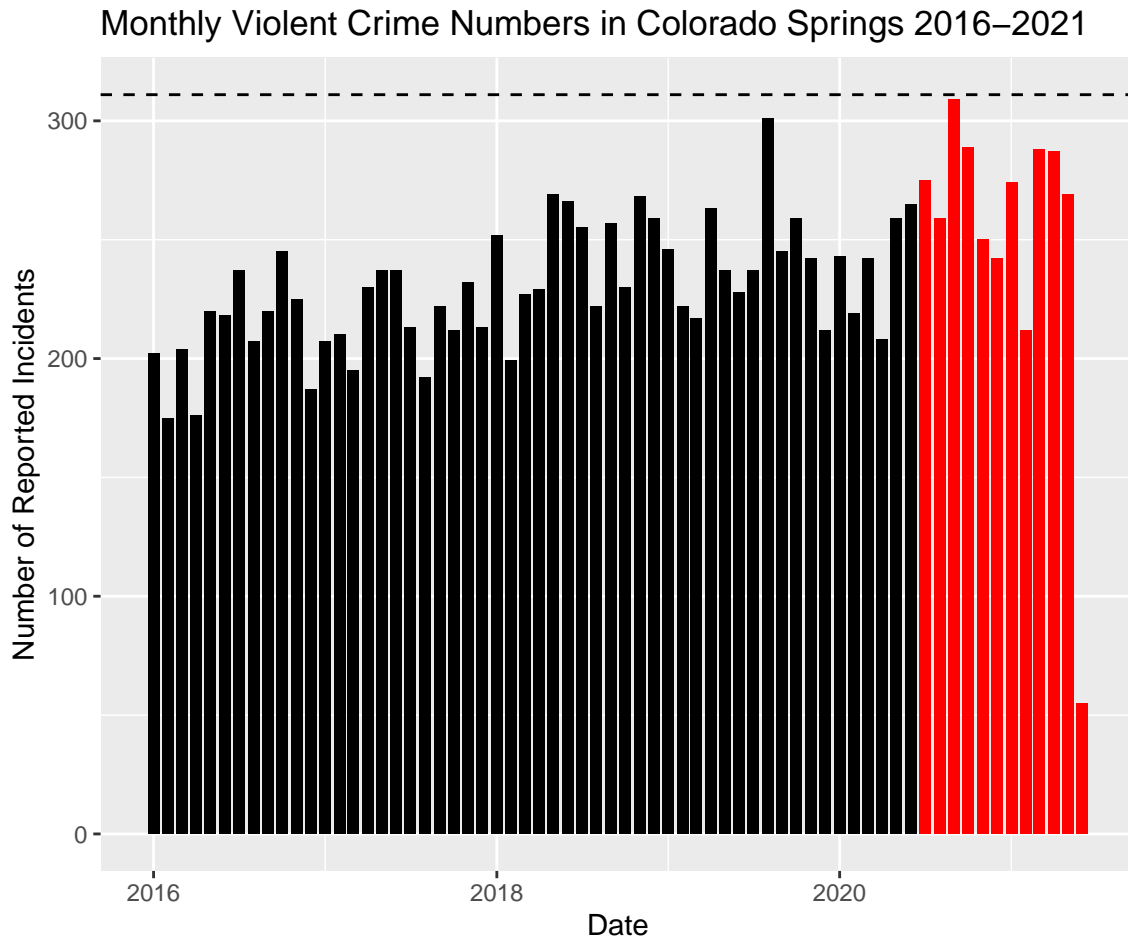
Similarly to Denver, Colorado Springs may have also experienced some increase in violent crime, although only one month after the passage of the police accountability bill surpassed the previous record for the highest number of violent crimes in a single month.

Monthly Property Crime Numbers in Colorado Springs 2016–2021
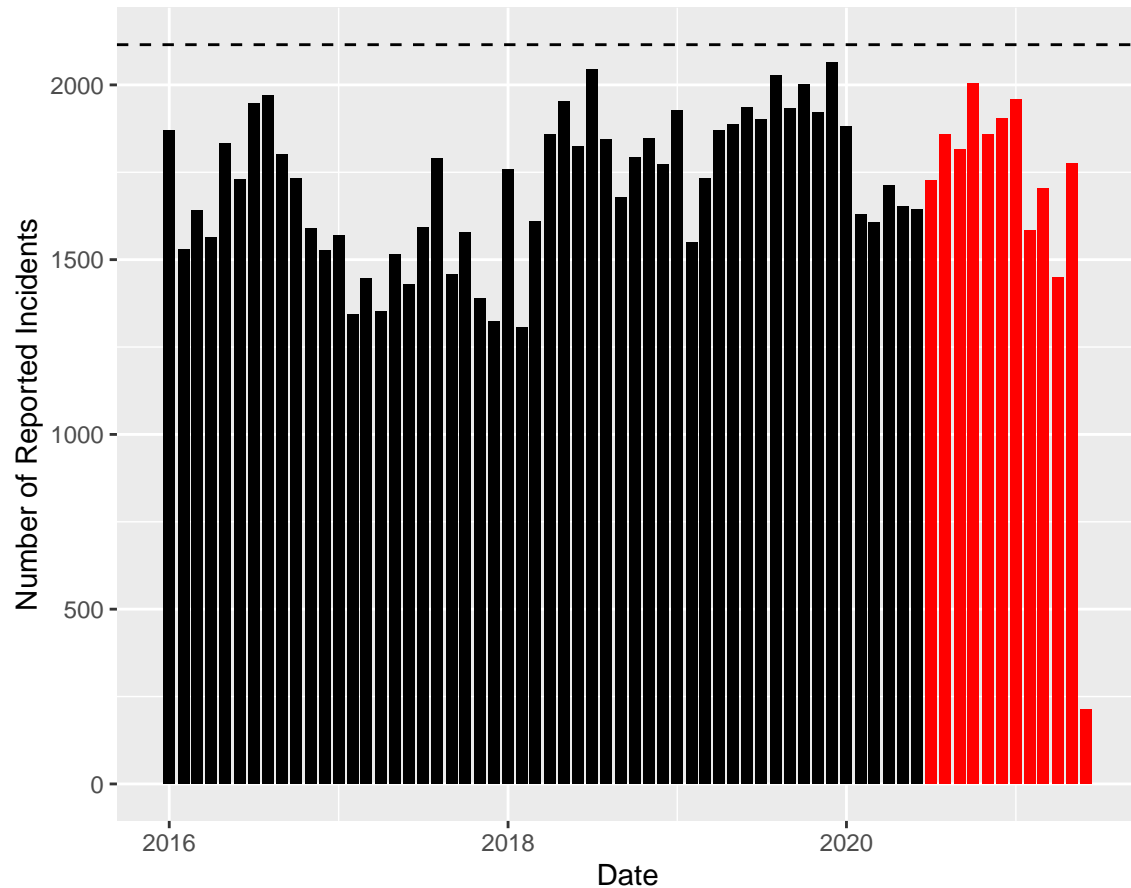
Diverging from Denver trends, however, Colorado Springs lacked the large rise in property crime numbers that Denver saw after the passage of qualified immunity. In fact, Colorado Springs property crimes may even have decreased over the same time period. The above bar graph demonstrates the possibility that other factors specific to Denver (such as worse recessionary effects from COVID or more distrust of police offiers) may have led to (or at least exacerbated) the major increase in property crimes, as opposed to the police accountability bill having an effect.

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) to determine if Denver's trends diverged from similar jurisdictions' crime trends.

Denver vs Control Monthly Violent Crime Numbers
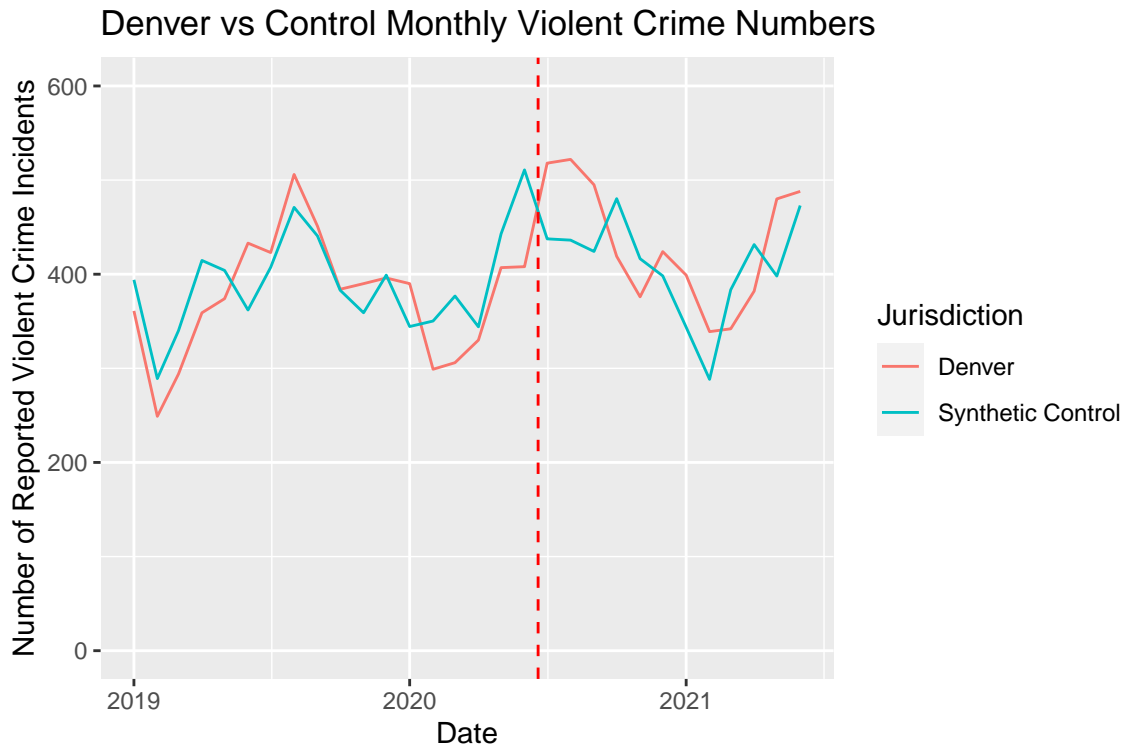
Denver's monthly violent crime incidents did not appear to vary significantly from similar jurisdictions. Although Denver experienced some increases in the summer immediately following the passage of the police accountability bill as well as some increases in the beginning of the 2021 summer not fully matched by the synthetic control, Denver's violent crime numbers one the whole appeared to closely follow many of the synthetic control numbers.

Denver vs Control Monthly Property Crime Numbers
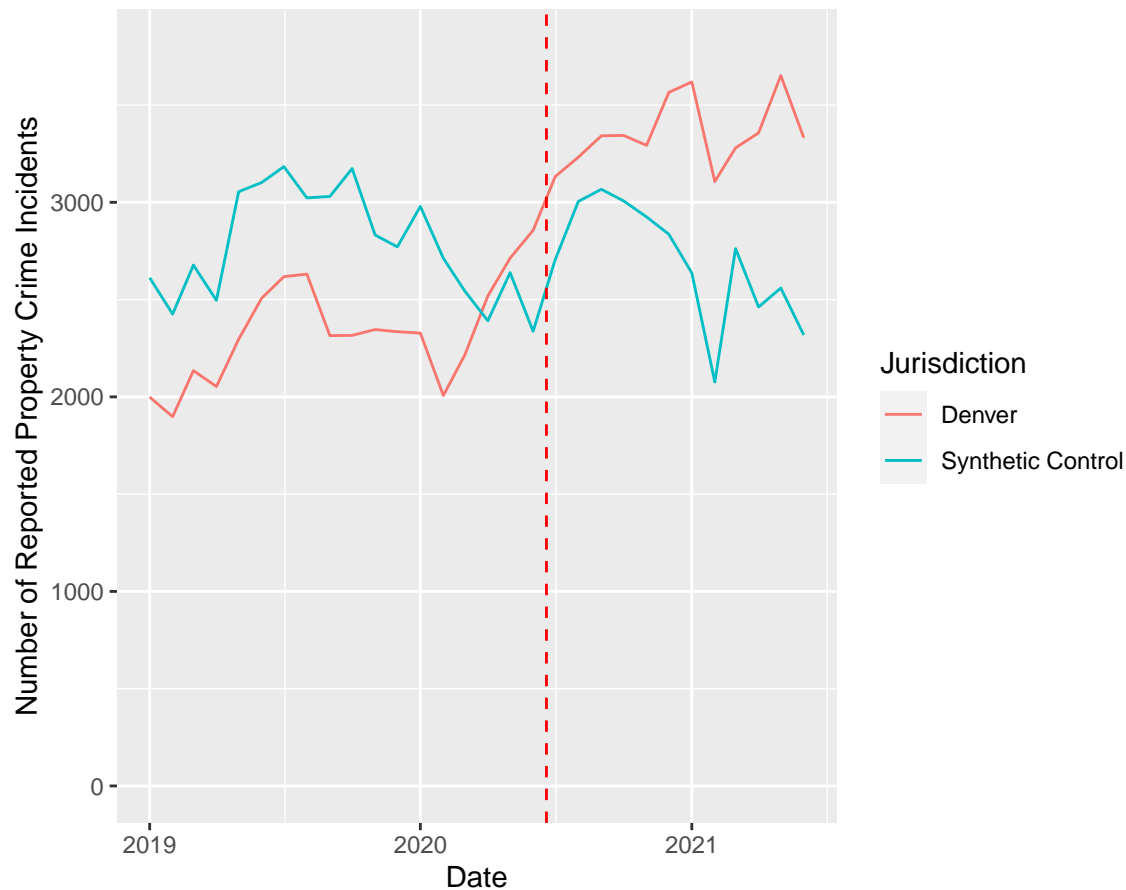
On the other hand, Denver's increase in property crime numbers was significantly greater than increases in other jurisdictions. From roughly February 2020 to July 2020, Denver property crimes steadily increased, while synthetic control numbers remained stagnant. Denver property cfrimes also remained high even after the summer, maintaining its much higher position compared to the synthetic control even as late as June of 2021. However, the graph comparing property crime numbers between Denver and the synthetic control unit is also much weaker than the violent crime graph, with the two lines significantly deviating from each other even before the passage of qualified immunity reform. The weaknesses in the graph may speak to possible dissimilarities between the synthetic control for Denver property crimes and true Denver.

Finally, we took the daily differences between the synthetic control and Denver between the two periods (the differences that we would later simulate on) and graphed the daily proportional increases in violent and property crimes.

Very little apparent increase in violent crimes in both Denver and control
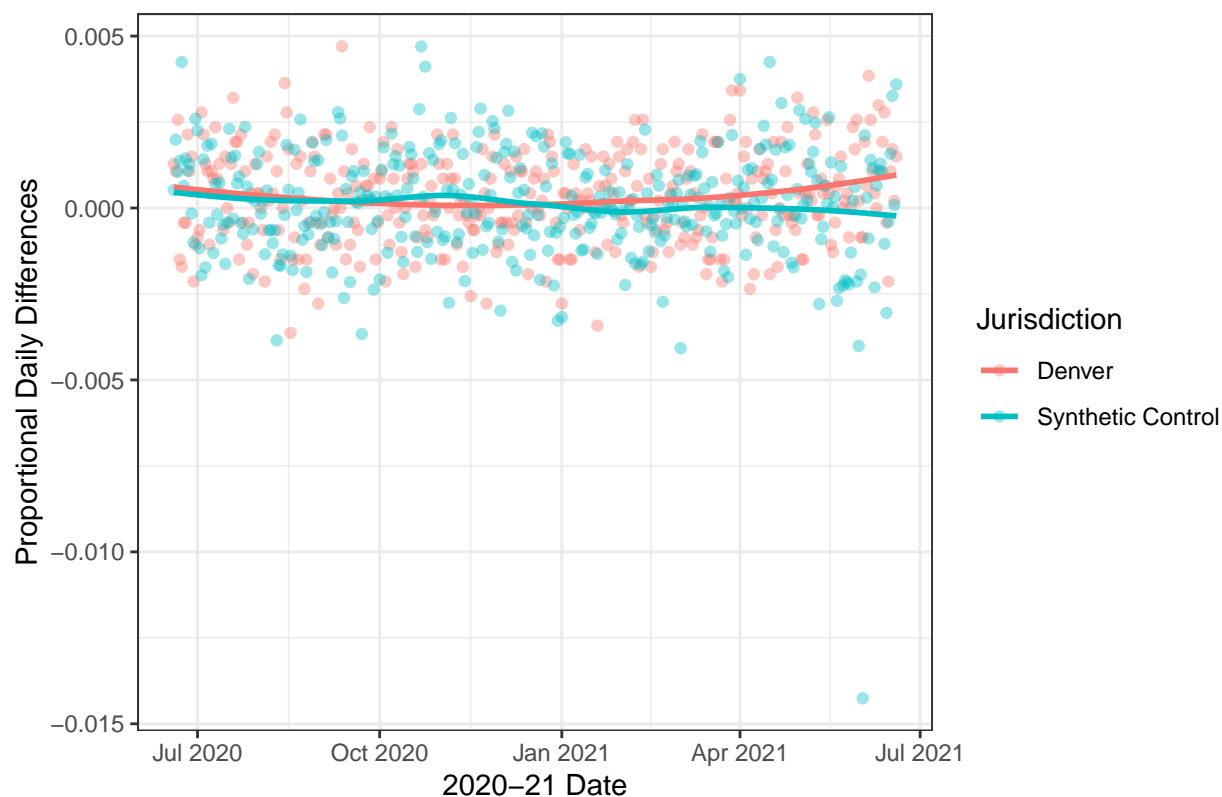
The graph does not appear to depict any significant daliy increase in violent crimes in either Denver or the synthetic control unit. The points are relatively randomly scattered and the trendlines are both roughly at 0 except for some divergence as the summer of 2021 approached. The graph also indicates the existence of a possible outlier on June 2, 2021. Upon further examination, the point at June 2, 2021 is mostly explainable by an extremely large amount of crime in Fort Smith in June 1, 2020 (the dates don't exactly match due to the extra leap day in 2019-2020) followed by 0 crimes in June 2, 2021. Since the difference was much below any other difference values, we decided to remove the June 2, 2021/June 1, 2020 data from the analysis to prevent means from being skewed.

Daily increases in Denver property crimes are greater than control incre
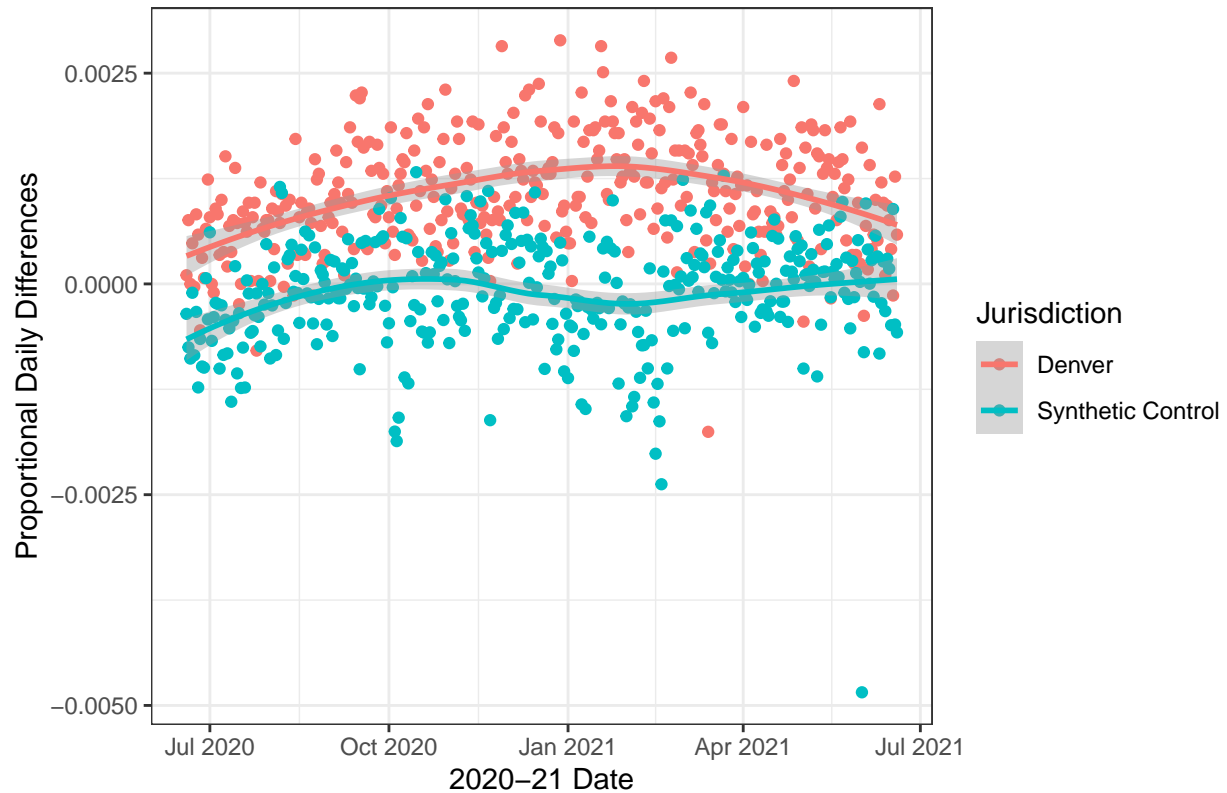
On the other hand, daily property crimes in Denver significantly increased, while the synthetic control trendline remained roughly at 0. Thus, even when comparing synthetic control differences with Denver, Denver still experienced a much larger increase in property crimes.

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.
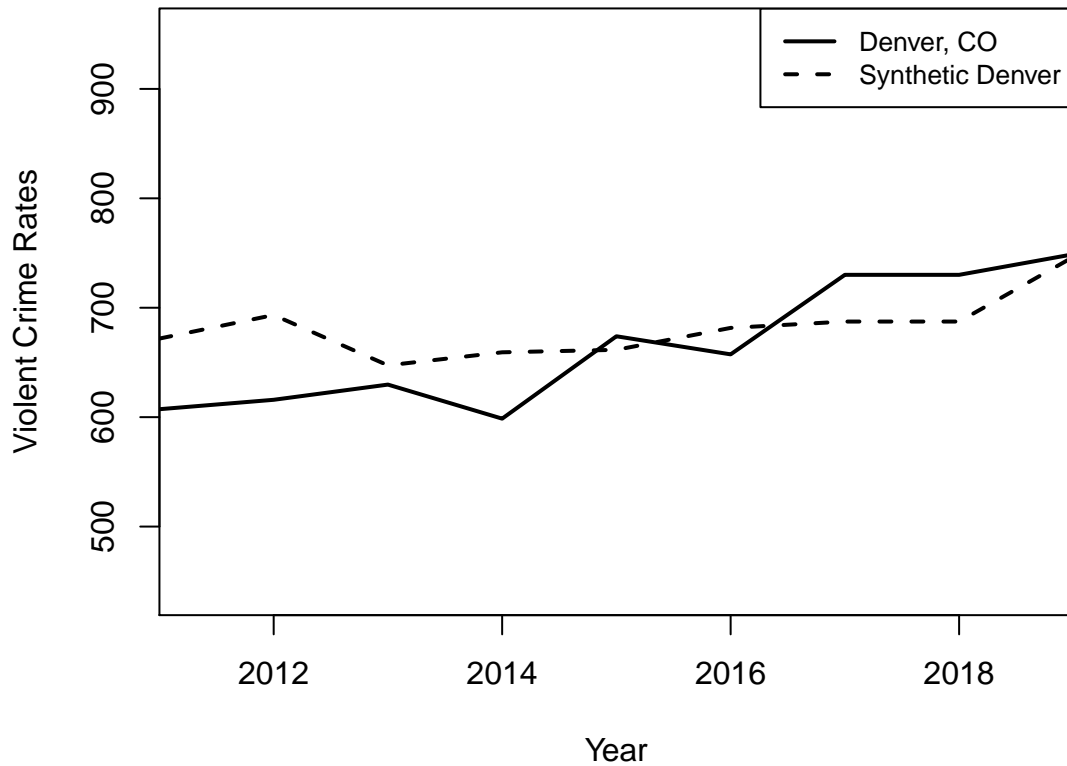
The first graph comparing the Denver annual violent crime rates with the synthetic control model's is relatively strong, with the synthetic control model's violent crime rates roughly matching some of the trends and the level of Denver's violent crime rate pre-treatment. There are some imperfections; 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:
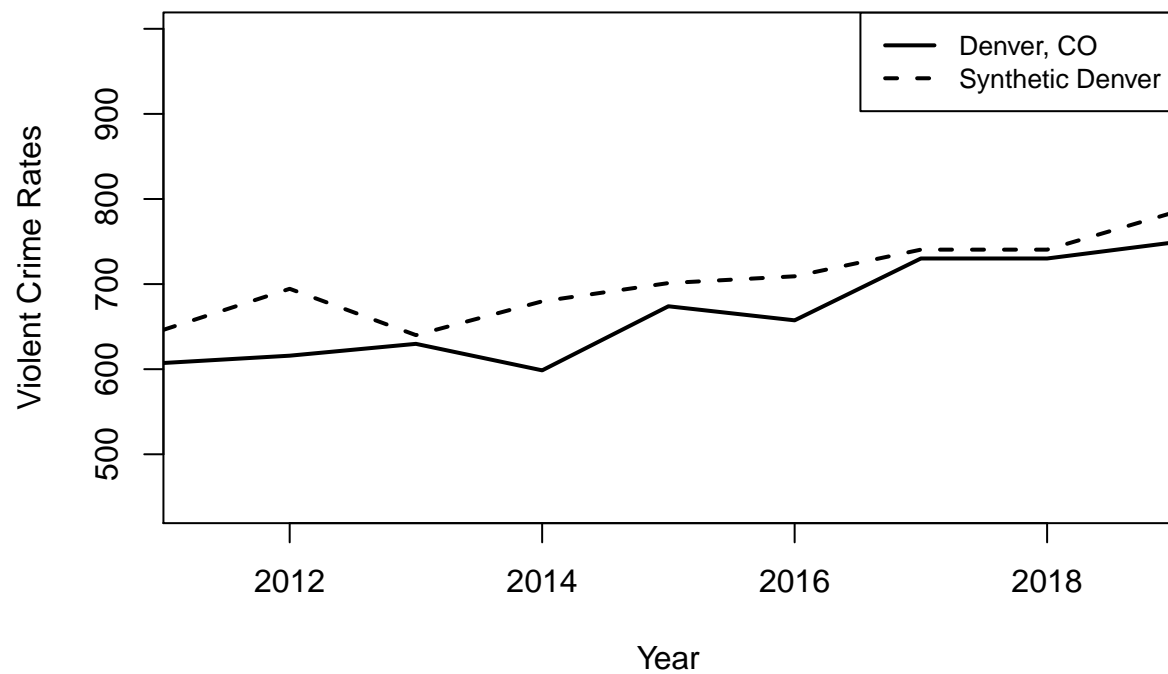
The synthetic of the 4 cities is slightly higher in terms of level but still follows the trend of true pre-treatment Denver roughly correctly. Thus, we determined the synthetic control to be similar enough to the real Denver to safely proceed with statistical testing.
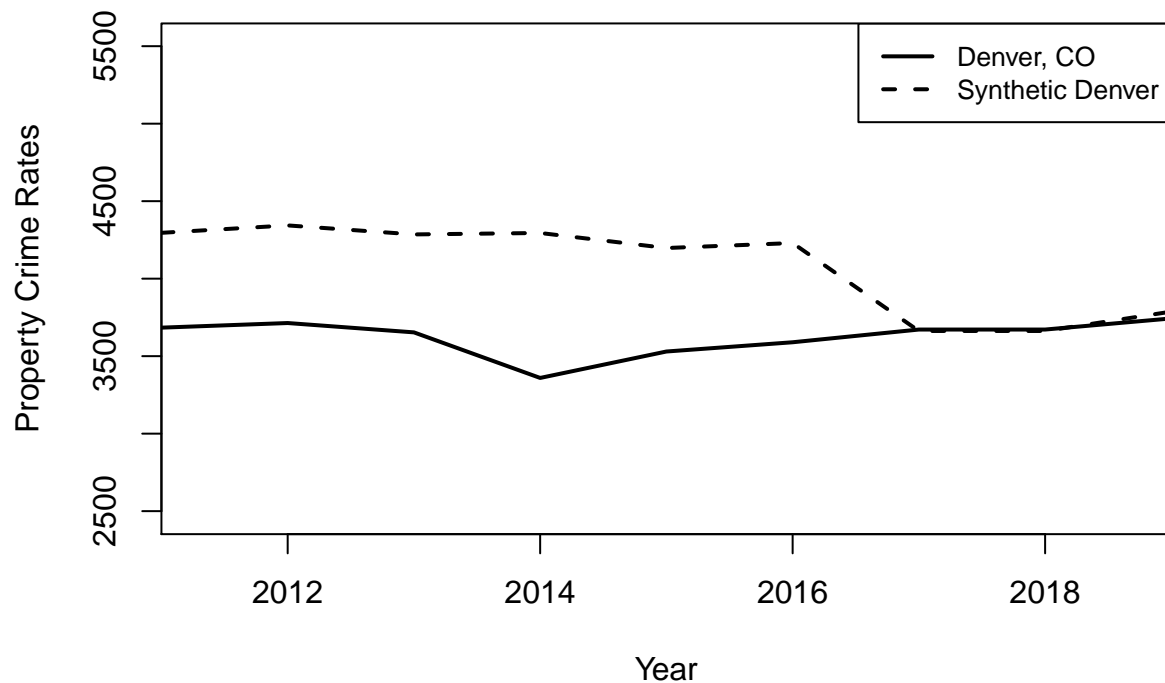
On the other hand, the plot comparing pre-treatment Denver and the synthetic control model for Denver indicates that the synthetic control model does not appear to be on the same level prior to 2017. From 2017 to 2019, the plot is roughly correct in comparing property crime rates, but the significant 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.
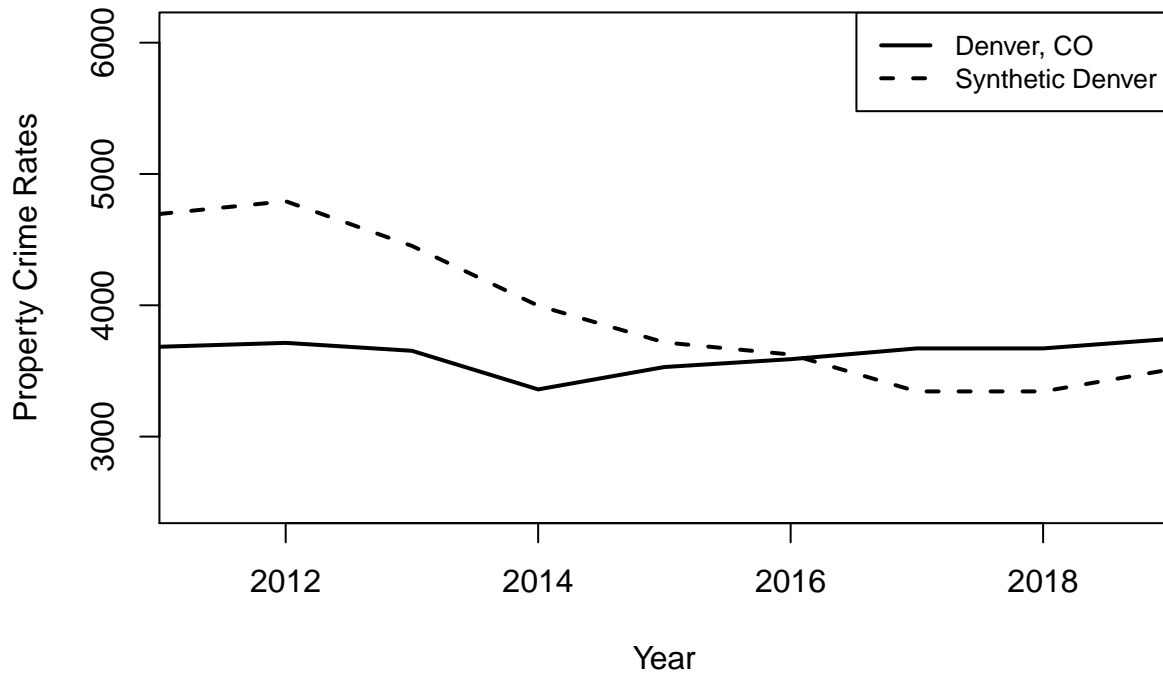
The synthetic of the top 5 cities of the previous synthetic control model is still significantly off in terms of following the trends of Denver in terms of property crimes. This may reflect crucial limitations in the ability of the synthetic control model to account for Denver property crime trends.

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 would 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 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

$H_0 : \mu_{Denver} = \mu_{Synthetic}$. 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.

$\alpha = 0.01$

Although we have population-level data, we utilized a bootstrapped simulation and hypothesis testing to determine if the difference between the Denver increases and the synthetic control increases could have resulted purely from chance.
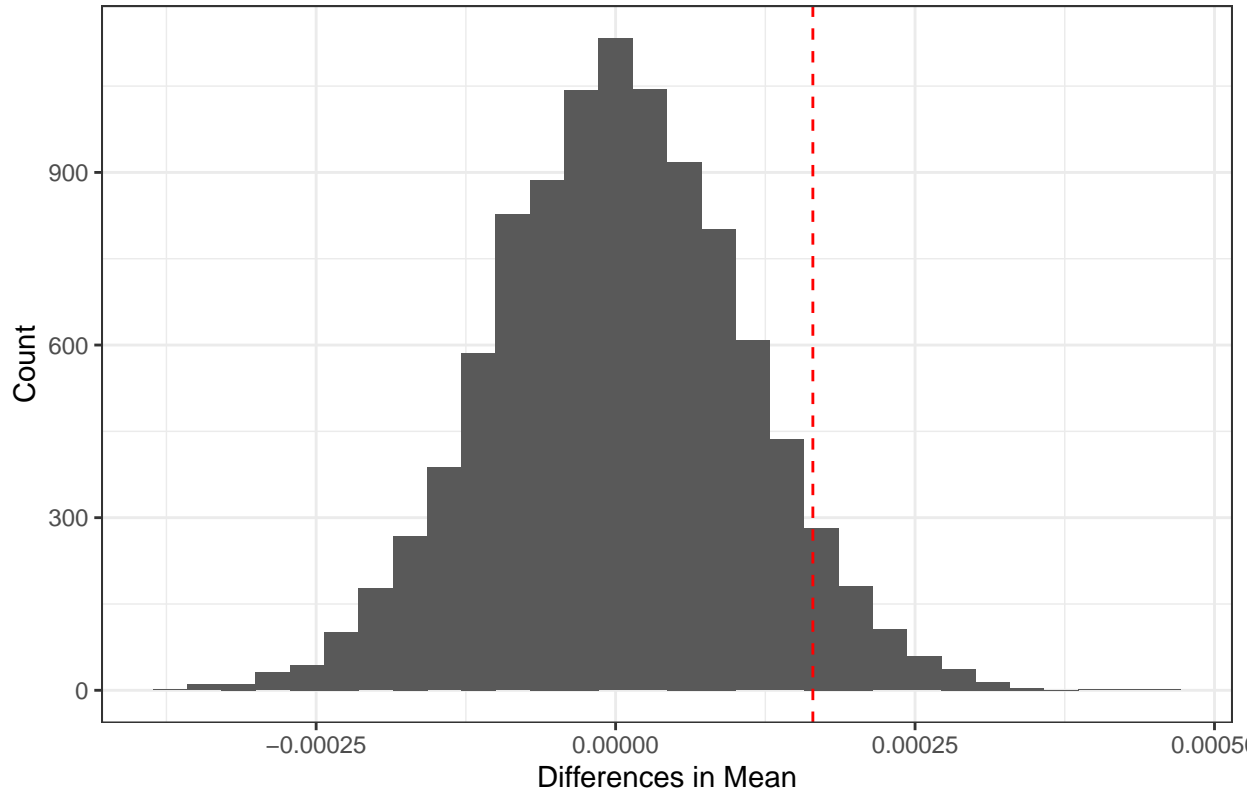
```
## # A tibble: 2 x 2
##   jurisdiction          mean
##   <chr>                <dbl>
## 1 Denver            0.000313
## 2 Synthetic Control 0.000148
```

```
## [1] 0.0593
```

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:

Null Distribution of Denver Violent Crime Differences in Mean

Because the p-value of 0.0593 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% or 5% 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. However, the data does provide sufficient evidence at the 10% level that Denver's average daily increase in violent crime after the passage of the police accountability bill is significantly greater than the synthetic control's average daily increase in violent crime.

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.

```
## # A tibble: 4 x 5
##   term          estimate std.error statistic  p.value
##   <chr>            <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)      391.       13.9     28.1   1.10e-34
## 2 time1             18.5      22.0      0.841 4.04e- 1
## 3 treated1         -15.2      19.7     -0.771 4.44e- 1
## 4 time1:treated1    37.9      31.1      1.22  2.28e- 1
```

The interaction variable for time and treated reflects the difference in difference estimate. Because the p-value of 0.277 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.
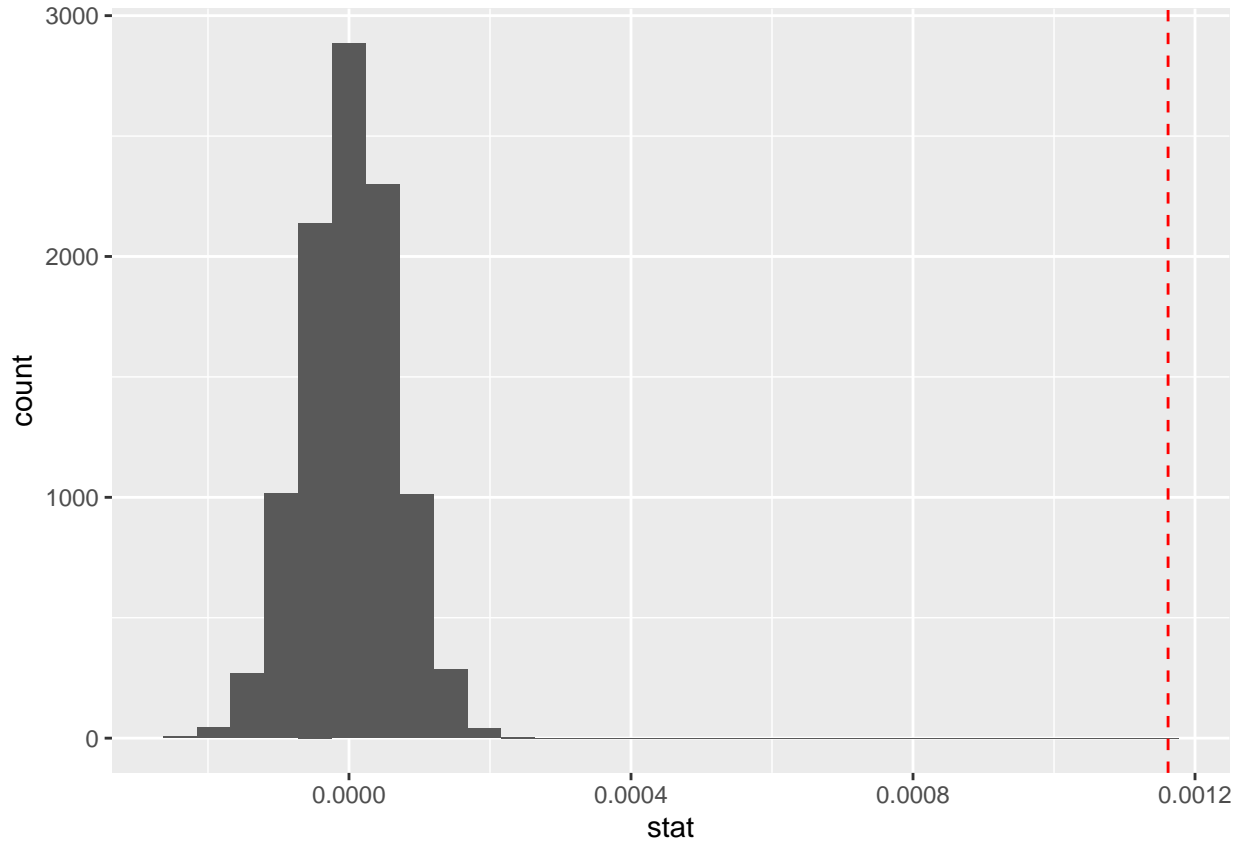
Because the p-value of 0 is far less than a reasonable alpha level of 0.01, we reject the null hypothesis. The data provides sufficient evidence to indicate that Denver's average daily proportional increase in property crime after the passage of qualified immunity was significantly greater than control 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.

Results and Discussion

The data does not provide evidence to indicate that Denver's increase in violent crime after qualified immunity significantly exceeded the violent crime increase in control cities. However, the data does suggest that Denver's increase in property crime did exceed the property crime increase in similar cities without qualified immunity reform. Unfortunately, because of the critical limitations in our data (see Methodology), we are unable to state that qualified immunity reform *caused* an increase in property crime in Denver. In addition, our exploratory data analysis may even provide evidence to the contrary; although Denver experienced a large rise in property crimes, Colorado Springs (the second largest jurisdiction in Colorado) did not experience much of a rise, making it possible that Denver-specific dynamics caused the increase in property crime. Once data from the census is released, we urge the conducting of more research to explore possible factors that may explain the rise in property crime in Denver that did not seem to be experienced at the same level elsewhere, even in the same state.

Our analysis is not over. In the future, we plan to take several steps to arrive at a stronger conclusion about whether qualified immunity reform had any clear effects on crime rates or police conduct. More immediately, we are creating a synthetic control for Colorado Springs to determine if Colorado Springs' increase in property crimes also exceeded control cities' increase. In addition, we also wish to eventually incorporate other large

jurisdictions within Colorado, such as Boulder or Aurora, into the analysis to determine if Denver's crime trend or Colorado Springs' crime trend or a third crime trend entirely can best explain crimes within the rest of Colorado. For instance, it is fully possible that Colorado Springs' lack of a property crime increase could be the outlier, while the rest of Colorado experienced an increase in property crime. In that case, the data would provide more support for the conclusion that qualified immunity reform could cause increases in property crime. In the long-term, we also expect to incorporate 2020 census data (as well as data on the COVID recession, police trust, and police employment) into the analysis as predictor variables to help statistically control for alternative explanations.