National Police Accountability Project
A Project of the National Lawyers Guild

SUPPORT HB 430 – Police Immunity and Accountability Act
Written Testimony of National Lawyers Guild-National Police Accountability Project, Keisha James, Staff Attorney
House Judiciary Committee – Tuesday, February 21, 2023

Dear Members of the House Judiciary Committee,

Thank you for the opportunity to provide comment on this important issue. The National Lawyers Guild National Police Accountability Project ("NPAP") is a nonprofit organization dedicated to holding law enforcement and corrections officers accountable to constitutional and professional standards. We strongly support the passage of HB 430, a bill that will enable victims of civil rights abuses to hold police officers accountable in state court without the shield of qualified immunity or statutory immunity. HB 430 will also increase accountability for police misconduct by allowing the Police Training and Standards Commission to review cases and revoke state certification.

The doctrine of qualified immunity has created a nearly insurmountable barrier for communities to hold police officers civilly liable in federal court for civil rights violations. Qualified immunity requires a victim of police misconduct to not only show that their constitutional rights were violated but prove that the violation was of "clearly established" law.[1] The Supreme Court has interpreted the "clearly established" law requirement to mean a plaintiff must be able to identify existing precedent that "squarely governs" the specific facts in their case in order to recover.[2]

There are many cases where an officer's patently unconstitutional conduct was shielded by qualified immunity because no prior defendant had been sued for similar behavior. For instance, in *Corbitt v. Vickers*,[3] a Georgia deputy sheriff accidentally shot a ten-year-old child lying on the ground while repeatedly attempting to shoot a pet dog that posed no threat. The circuit court held that the

[1] *Harlow v. Fitzgerald*, 457 U.S. 800, 818 (1982).
[2] *Kisela v. Hughes*, 138 S. Ct. 1148, 1153 (2018).
[3] 929 F.3d 1304 (11th Cir. 2019).

deputy was entitled to qualified immunity because there was no prior case with the same unique set of facts. There are dozens of other equally ludicrous and unjust outcomes that have resulted from the doctrine of qualified immunity.[4]

While the Maryland legislature cannot eliminate qualified immunity in federal courts, it can provide people in this state with an alternative method to vindicate their rights in state court. Although Maryland provides a common law private right of action that allows people to recover for violations perpetrated by local police officers,[5] state police officers are still granted immunity by statute and, as such, victims of civil rights abuses do not currently have a clear path to hold state police officers accountable.[6] In *Smith v. Md. State Police Dep't*,[7] Maryland State Police officers stopped a car without probable cause and commanded a K-9 to attack an individual after he had already been restrained. The court found that the officers were not entitled to qualified immunity because prior cases had clearly established the constitutional rights at issue. If these prior cases had not existed, however, the officers would have been shielded by qualified immunity. HB 430 would prevent police officers who deprive people of their constitutional rights from invoking qualified immunity as a defense to liability, enabling victims to vindicate their rights whether or not there are prior cases with similar facts.

Although Maryland passed reforms two years ago that addressed various areas of policing, qualified immunity remains unaddressed. The repeal of the Law Enforcement Officers' Bill of Rights ("LEOBOR") was a key step in increasing transparency in policing and ensuring that police officers face employment consequences for misconduct. However, LEOBOR's repeal does not help victims of civil rights abuses who have suffered tangible injuries as a result of police misconduct. HB 430 seeks to help these victims. It is important to note that Maryland would not be the first state to eliminate strict immunity defenses for

---

[4] *See* National Police Accountability Project, Expanding Pathways to Accountability: State Legislative Options to Remove the Barrier of Qualified Immunity, available at https://engage.nlg-npap.org/sites/default/files/2022-06/Qualified-Immunity-White-Paper-Final.pdf.
[5] *See Ritchie v. Donnelly*, 597 A.2d 432 (Md. 1991); *Clea v. Mayor and City Council of Baltimore*, 541 A.2d 1303, 1312 (Md. 1988).
[6] *See* Md. Code Ann., Cts. & Jud. Proc. § 5-522; *Lee v. Cline*, 863 A.2d 297 (Md. 2004).
[7] Civil Action No. GLR-18-2836, 2019 U.S. Dist. LEXIS 170767 (D. Md. Sep. 30, 2019).

state court civil rights actions against police officers. Montana,[8] Colorado,[9] and New Mexico[10] have all rejected qualified immunity defenses for state constitutional actions against law enforcement officers. Most recently, the Nevada Supreme Court ruled that qualified immunity is not a defense in cases where government officials conduct unlawful searches and seizures.[11]

The efforts in other states to end qualified immunity were met with the same general questions posed by those in opposition to HB 430. But these questions stem from a misunderstanding about how qualified immunity functions in practice:

- **Will eliminating qualified immunity cost the government more money?**
    - **No.** An increased risk of liability for violating civil rights will help save law enforcement agencies money in the long run by incentivizing them to deter bad behavior through strengthening policies and procedures and improving training. If officers refuse to follow department policies, procedures, and training, agencies will be compelled to discipline and terminate those officers rather than risk liability for their misconduct. Moreover, forcing Maryland residents to contend with statutory and qualified immunity does not save costs but shifts them from law enforcement officers and agencies to victims of police misconduct. Victims of police brutality in particular experience tangible consequences, including medical costs, lost wages, and emotional trauma. Eliminating qualified immunity will enable victims to be compensated for their injuries and incentivize agencies to prevent misconduct in the future.

[8] *Dorwat v. Caraway*, 58 P.3d 128, 131 (Mont. 2002).
[9] COLO. REV. STAT. ANN. § 13-21-131 (2020).
[10] New Mexico Civil Rights Act, 2021Bill Text NM HB 4.
[11] Nick Sibilla, *Nevada Supreme Court Upholds The Right To Sue The Government, Blocks Qualified Immunity*, Forbes (Jan. 12, 2023), https://www.forbes.com/sites/nicksibilla/2023/01/12/nevada-supreme-court-upholds-the-right-to-sue-the-government-blocks-qualified-immunity/?sh=4e992837588e.

- **Will eliminating qualified immunity lead to a significant increase in crime because officers will be afraid to do their job?**
    - **No.** Officers who are making reasonable, good faith decisions when carrying out their duties and following their training and department policies do not need qualified immunity to continue doing their jobs—they are already protected by the Fourth Amendment. An examination of data from Denver and Colorado Springs shows that violent crime rates have remained the same since qualified immunity reform was passed.[12] Crime rates in Denver also remained consistent with cities with similar populations and demographics.[13]

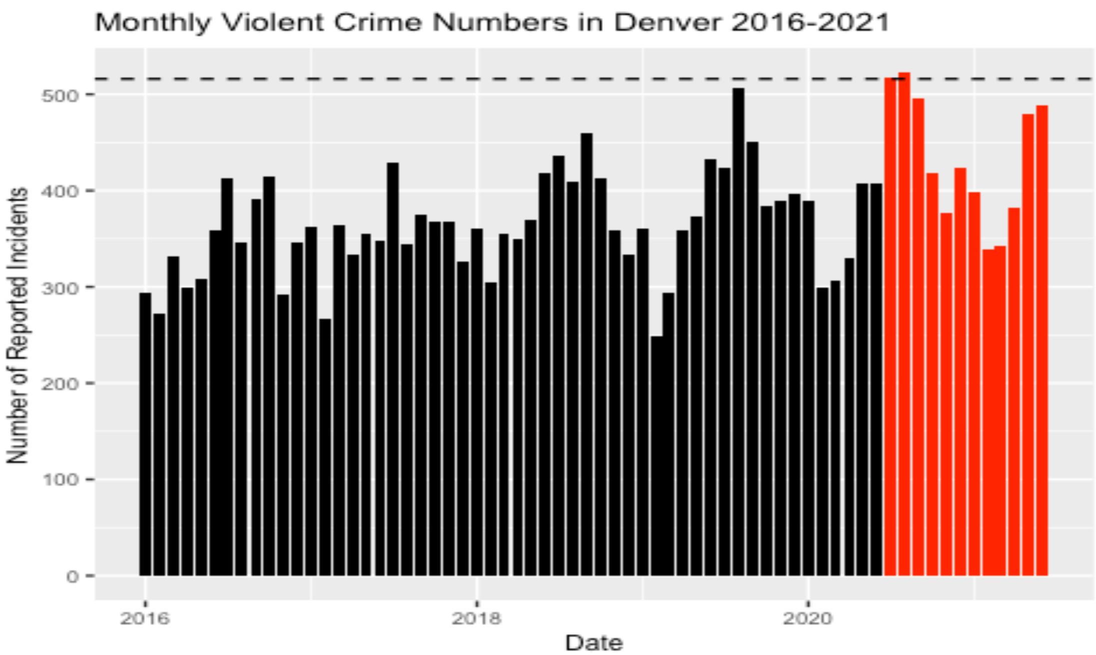
- **Will eliminating qualified immunity result in court inefficiencies?**
    - **No.** Litigating the issue of qualified immunity is not only costly for the parties but prolongs cases in front of the court.[14] Following the elimination of qualified immunity in New Mexico, civil rights attorneys reported a decrease in litigation delays.[15] Eliminating qualified immunity as a defense would streamline cases.

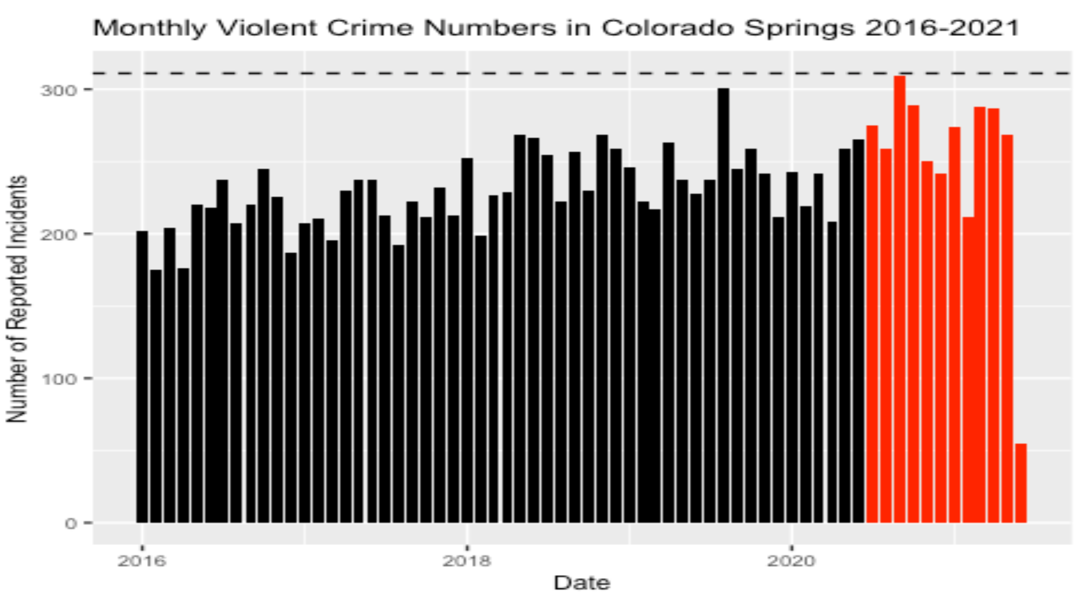
We strongly urge you to make a favorable report on HB 430. Thank you, again, for the opportunity to provide comment on this important issue.

---

[12] *See* Exhibit A (attached).
[13] *See* Exhibit B (attached); *see also* Exhibit C (attached).
[14] Joanna C. Schwartz, *Qualified Immunity's Selection Effects*, 114 Nw. U. L. Rev. 1101, 1119 (2020) ("[Q]ualified immunity doctrine increases the cost, time, and complexity of litigating police misconduct cases.").
[15] *See* Exhibit D (attached).

Exhibit A



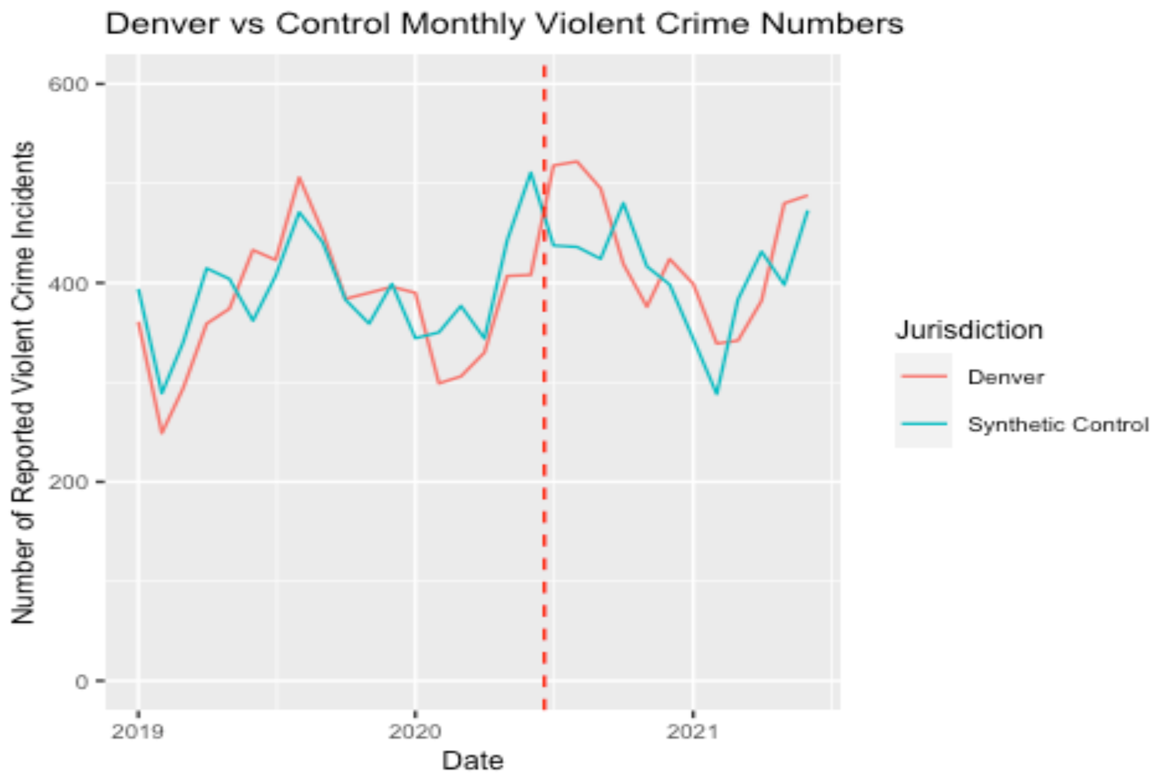


Monthly Violent Crime Numbers 2016-2021: Each bar represents the number of reported incidents in a single month. Red bars represent the months following the passage of the police accountability legislation in June 19, 2020. Data from June 20-30 is included in the month immediately preceding the red bars (June 2020). The black dotted line represents 10 reported violent incidents above the previous maximum number of offenses in a single month in the four years prior to legislation. Data was derived from the FBI,[1] CSPD,[2] and DPD.[3]

---

[1]https://crime-data-explorer.app.cloud.gov/pages/explorer/crime/crime-trend
[2] https://policedata.coloradosprings.gov/Crime/Crimes-Against-People/ghs7-nqyk.
[3] https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-crime.

Exhibit B



Monthly Violent Crime Numbers in Denver Compared to Control 2019-2021: The red dashed line represents the passage of the police accountability legislation in June 19, 2020. The control continues to mostly track Denver violent crimes even after the passage of the police accountability law. The synthetic control model was created by culling data from American municipalities from 2011-2019 that had both UCR and American Community Survey census data. From there, the cities were filtered to only include cities greater than 50000 in population to identify roughly 500 similar jurisdictions to Denver. "Similar" cities were identified by analyzing the following factors: population size, single female-led family household percentage, high school graduate or higher percentage, 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.

Exhibit C

Statistical Report on Colorado's Qualified Immunity Reform and Crime Rates

By Andrew Qin and the National Police Accountability Project

Abstract: We wished to determine if Colorado's police accountability reform (SB 217) could have caused a significant increase in violent and property crime rates in Colorado's most populous jurisdictions. We compiled a database of 88 comparable jurisdictions and ran synthetic control models for Denver, Colorado Springs, and Aurora to determine if the jurisdictions' crime rates were greater than expected. We then checked if the difference between the synthetic controls and the jurisdictions were statistically significant through placebo testing. We ended up finding that 1) Colorado's three most populous jurisdictions did not experience significantly higher violent crime rates in 2020 and 2021 (post-treatment) compared to their controls after placebo testing, and 2) The Denver-Aurora MSA did experience significantly higher property crime rates in 2020 and 2021 compared to their controls, but Colorado Springs did not experience a significant increase in property crime rates. We concluded that the data does not provide evidence to indicate that the statewide police accountability reform caused a property crime or violent crime increase.

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. 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 argue that qualified immunity prevents police officers from being held accountable for misconduct in civil litigation. Supporters of qualified immunity argue that efforts to limit the doctrine would prevent police officers from effectively performing their jobs for fear of frivolous lawsuits. They suggest that eliminating qualified immunity will therefore indirectly lead to a rise in crime. In this statistical report, we aim to provide preliminary data-driven insights on the effects of recently passed qualified immunity legislation on violent and property crime rates in major urban jurisdictions.

On June 19, 2020, Colorado became the first state to implement qualified immunity reform as part of the omnibus Enhance Law Enforcement Integrity Act (Senate Bill 20-217). The Act prevented officers from using qualified immunity as a defense against civil liability for violations of constitutional rights. The Act also included several other measures, including a ban on chokeholds, harsher penalties for illegal use of force, mandatory internal reporting, narrower use of force guidelines, and new guidelines on acceptable responses to protests. Most reforms took effect immediately (including the qualified immunity provisions) or on September 1, 2020. Other reforms, including data collection and body camera requirements, will take effect in 2022 and 2023. Three other major jurisdictions have passed measures to limit or reform qualified immunity. In August of 2020, Connecticut passed a measure that limited qualified immunity but only took effect in July of 2021. New Mexico and New York City passed measures to limit qualified immunity in 2021.

Because Colorado was the first jurisdiction to implement qualified immunity reform, we analyzed Colorado data to determine the plausible effects of reform on crime rates. We asked the following question: Was the passage of qualified immunity reform in Colorado in June 2020 correlated with significant increases in violent and property crime rates compared to increases in control jurisdictions? We further narrowed the scope of our analysis to Denver, Colorado Springs, and Aurora—the three largest jurisdictions within Colorado – due to missing 2021 statewide data on crime rates.

We should note, however, that although we wished to determine the effects of qualified immunity reform on crime rates, we could not disentangle the effects of the other reforms in the police accountability bill. As we discussed above, SB 217 was an omnibus police reform bill, with several measures enhancing police accountability. If we observed any statewide causal effect, our analysis could not differentiate between which measure resulted in the effect. We believe there is some possibility that the effects of other elements of the police accountability law could have been partly controlled for by coincidence. The reforms in SB 217 outside of qualified immunity and civil liability reform are shared with several other jurisdictions; 17 states passed similar bans on chokeholds, and 30 states passed some form of police accountability law. Additionally, we only excluded the jurisdictions that passed qualified immunity reform from our control set. Jurisdictions that passed police accountability laws matching Colorado's in every way except for qualified immunity reform were included in the analysis. Nonetheless, because we could not fully control for the other reforms passed in the law, we will only discuss our conclusions in the context of the police accountability law more broadly.

Although we attempt to establish some level of causation in this study by using a synthetic control method, we lack the volume of observational data needed to successfully conduct causal inference. In particular, we lack observations on key lurking variables, including the effects of the COVID-19 pandemic on poverty rates in each jurisdiction, 2020 and 2021 census data, and shifting community attitudes towards policing.[1] Much of this data will only be released a few years from now, limiting the contours of the present analysis. However, due to the prescience of the qualified immunity question, we decided to produce this preliminary report to at least illustrate the plausible effects of SB 217 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.

Overview of Available Data

Many of our Methodology choices can only be understood in the context of the available data and policy context at our disposal. We obtained three different types of data from three sources, almost entirely official (with the exception of land area data, which was obtained from a website reporting census data).

First, we received socioeconomic indicator data from the 2011-2019 American Community Survey Five-Year estimates as found through the census data website. The predictor data we utilized was organized "by Place," meaning the data was largely aggregated in terms of local unit boundaries (towns, cities, census-designated places ("CDPs"), etc.). This data was collected with the intention of serving as crime predictor data. Unfortunately, at the time of the creation of this report, neither 2020 nor 2021 census data had been publicly released.

Second, we collected crime rate data by state and city through the Uniform Crime Reports ("UCR") released by the FBI. This data was complete from 2011-2020, and the first three quarters of both 2020 and 2021 had been released in the Quarterly UCR from roughly 155 large agencies (limited by the number of agencies that reported their crime rates). This data included statistics on jurisdiction population, violent crime numbers, property crime numbers, and numbers for individual crimes (such as forcible rape, nonnegligent homicide/murder, larceny, etc.). The quarterly data was not disaggregated by quarter.

Third, we received incident level crime data by downloading the data from various agency websites and submitting FOIA requests for agencies that had not released their data publicly (such as Champaign Police Department). Several times, these FOIA requests returned data unfeasible to work with (such as PDF reports of individual crimes), were deemed too costly (totaling greater than $100 for smaller agencies), or were flatly denied on the basis that data was not kept or that state FOIA laws only permitted in-state residents to make FOIA requests (in the case of Clarksville). As a result, the usefulness of FOIA requested data was limited; however, we incorporated the data that we could obtain using this method into Methodology B analysis.

[1] By "community attitudes towards policing," we refer to the possibility that increasing distrust of police officers may have changed citizens' perspectives on crime and cooperation with police, both recognized by the FBI as variables affecting crime.

In sum, the data that we could obtain was limited in scope, largely due to the combination of limited fiscal/temporal resources, difficulty in obtaining data from police departments, and late releases of census and UCR data. The data limitations then harmed the soundness of the analysis. Nonetheless, we managed to obtain enough data to derive meaningful insights on crime rates in treated and control jurisdictions.

Methodology

In this study, we employed two different methodologies, one to incorporate crime rates from all jurisdictions for 2020-2021 ("Methodology A") and the other to increase the accuracy of the treatment date ("Methodology B"). We have primarily incorporated findings from Methodology A, but findings from Methodology B (performed months before Methodology A was performed) are included in the appendices of this report.

In both methodologies, we employed, to varying extents, a synthetic control methodology as described by Alberto 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 creating control jurisdictions based on average predictor numbers over time, removing the effects of researcher bias from the analysis.

Sampling Methodology

Using data provided by the UCR, we created a dataset of 93 urban jurisdictions and tracked their violent and property crime numbers from 2011 to the first three quarters of 2021. Jurisdictions were chosen on the basis of two criteria: First, the jurisdictions' violent and property crime numbers must have been published every year (from 2011 to 2021) by the UCR. Because the UCR's 2021 quarterly crime report only published figures from large self-reporting jurisdictions, the sample of jurisdictions is influenced by self-selection sampling bias; those jurisdictions that chose not to report figures for one of the years are automatically excluded. Second, jurisdictions must have been greater than 85,000 in population according to UCR estimates for every year from 2011 to 2020 (no population data for 2021). This measure is meant to exclude excessively small jurisdictions at the beginning time period that experienced extreme population growth. The number 85,000 was chosen, in part, as a value for a jurisdiction that could experience average population growth from 2011 to reach at least 100,000 by 2020. The second criterion only excludes four small jurisdictions from the analysis, each of which likely did not match the dynamics of larger urban jurisdictions like Colorado Springs and Denver.

After creating the database of crime numbers, we then compiled a set of yearly socioeconomic indicators from each jurisdiction to serve as predictor values for the 'Synth' package to average when creating a synthetic control. We chose indicators on the basis of sociological evidence that such indicators could serve as moderately strong predictors of metropolitan violent or property crime rates. Based on the results found in Wells and Weishelt's "Explaining Crime in Metropolitan and Non-Metropolitan Counties," we chose to record jurisdictions' high school education percentage, residential stability (or percent of people living at the same property that they lived at one year ago), percentage of population over 18, percentage of population who is white, percentage of population who is self-employed, unemployment rate, median income, child poverty rate, and population density (population divided by land area). Certain variables recorded in the Wells and Weishelt study were excluded from our analysis because they were either found to be largely non-significant for metropolitan counties in the study (such as South vs. non-South or owner-occupied housing), had missing data for some years (such as single female-led household percentage), or were difficult to collect (such as percentage of population that voted in the last election). These indicators were collected by jurisdiction for 2011 to 2019 from the census tables. If a jurisdiction was missing data from any of those years on any of the collected variables, the jurisdiction was excluded from the analysis.

Certain errors occurred when combining census data with UCR data, particularly around the naming schemes of the cities. While the UCR names cities by their given names, the census data often adds addendum names such as “CDP,” “city,” “town,” and others. We attempted to correct for these errors by erasing addendum words from census names (for instance, removing " City" from names as in the case of “Boise City, Idaho” or “Houston City, Texas”). For large jurisdictions (usually above 100,000 in population), we further went back and individually corrected names to match. We believe we caught most of these errors, but some errors inevitably slipped through the cracks, leading to randomly lost data. Regardless, we find it unlikely that these random errors significantly hindered our analysis.

Because we lacked predictor data for 2020 and 2021, we extrapolated predictor data from 2019 to 2020 and 2021. In other words, 2020 and 2021 predictor data (outside of population and population density) were equivalent to 2019 data. Additionally, 2021 population and population density were extrapolated from 2020 population figures. We do not argue that this extrapolation is a fair representation of reality; of course, with the COVID-19 pandemic and the George Floyd protests of 2020, socioeconomic indicators in 2020 will be different from those in 2019. Extrapolating skewed our pre-treatment predictor averages to some extent, but we do not think it invalidates our results. We further discuss the implications of this choice in the “Methodology A” section.

In the end, we had a dataset of 93 jurisdictions, with crime data from 2011 to 2021 and predictor data from 2011 to 2019. In total, our dataset had 1023 observations and 24 variables. In Table 1, we display the first 20 rows of our dataset.

Table 1: First 20 Rows of Dataset (split into 3 pages)

NAME	population	violent_crime	property_crime	year	violent_crime_rate	property_crime_rate
Alexandria, VA	141638	252	3181	2011	177.918	2245.866
Alexandria, VA	145892	243	2990	2012	166.562	2049.461
Alexandria, VA	148519	258	2967	2013	173.715	1997.724
Alexandria, VA	151065	276	2960	2014	182.703	1959.421
Alexandria, VA	152710	312	2854	2015	204.309	1868.902
Alexandria, VA	155319	286	2798	2016	184.137	1801.454
Alexandria, VA	158256	262	2482	2017	165.555	1568.345
Alexandria, VA	162588	260	2482	2018	159.913	1526.558
Alexandria, VA	162258	288	2517	2019	177.495	1551.233
Alexandria, VA	161525	295	2793	2020	182.634	1729.144
Alexandria, VA	161525	235	1783	2021	145.488	1103.854
Ann Arbor, MI	113848	261	2549	2011	229.253	2238.950

Ann Arbor, MI	115008	227	2726	2012	197.378	2370.270
Ann Arbor, MI	116799	247	2525	2013	211.474	2161.834
Ann Arbor, MI	117768	194	2200	2014	164.731	1868.080
Ann Arbor, MI	118730	228	2364	2015	192.032	1991.072
Ann Arbor, MI	117688	213	2051	2016	180.987	1742.744
Ann Arbor, MI	121930	259	2108	2017	212.417	1728.861
Ann Arbor, MI	122571	270	1932	2018	220.280	1576.229
Ann Arbor, MI	122893	309	2124	2019	251.438	1728.333

NAME	female_household[2]	hs	res_stability	over_18	white_percent	self_employed	unemployment	income
Alexandria, VA	8.6	91.0	78.0	83.0	54.3	4.5	4.5	82899
Alexandria, VA	8.6	91.7	78.2	82.9	53.6	4.5	5.1	83996
Alexandria, VA	7.9	91.2	78.3	82.7	53.1	4.4	5.0	85706
Alexandria, VA	8.1	91.3	78.2	82.5	52.7	4.7	4.7	87319
Alexandria, VA	8.6	91.5	77.2	82.3	52.4	4.7	4.5	89134
Alexandria, VA	8.6	91.4	77.4	82.0	52.0	4.7	4.0	89200
Alexandria, VA	8.3	91.4	78.0	81.9	51.8	4.9	3.9	93370
Alexandria, VA	8.7	92.5	78.3	81.7	51.8	4.9	3.3	96733
Alexandria, VA	3.9	93.0	79.0	81.8	51.9	4.8	3.0	100939
Alexandria, VA	3.9	93.0	79.0	81.8	51.9	4.8	3.0	100939
Alexandria, VA	3.9	93.0	79.0	81.8	51.9	4.8	3.0	100939
Ann Arbor, MI	6.8	96.5	64.1	85.4	69.9	4.7	7.3	53377
Ann Arbor, MI	6.8	96.5	64.4	85.8	69.8	4.7	7.2	53814
Ann Arbor, MI	6.7	96.5	65.2	85.5	69.8	4.3	7.6	55003
Ann Arbor, MI	6.6	96.4	64.6	85.6	69.1	4.3	7.1	56835
Ann Arbor, MI	6.6	96.4	64.0	86.0	68.9	4.3	6.5	55990
Ann Arbor, MI	6.4	96.8	64.1	86.1	68.7	4.2	5.6	57697
Ann Arbor, MI	6.0	96.8	64.5	86.6	68.6	4.2	5.3	61247
Ann Arbor, MI	6.0	97.1	64.6	86.9	67.4	4.3	4.6	63956
Ann Arbor, MI	2.6	97.3	66.1	87.2	67.5	4.1	3.9	65745

NAME	received_snap[3]	child_poverty	owner_occupied	land_area	pop_density	obs_num	id
Alexandria, VA	3.4	12.4	45.0	15	9442.533	43	1
Alexandria, VA	4.2	13.0	43.9	15	9726.133	44	1
Alexandria, VA	4.6	13.8	43.3	15	9901.267	45	1
Alexandria, VA	4.8	13.7	42.7	15	10071.000	46	1
Alexandria, VA	4.6	12.8	42.5	15	10180.667	47	1
Alexandria, VA	5.0	15.2	42.2	15	10354.600	48	1
Alexandria, VA	4.6	17.7	43.1	15	10550.400	49	1
Alexandria, VA	4.4	18.6	42.9	15	10839.200	50	1

[2] Originally, we tracked single female-led household percentage, but we soon found out that the 2019 ACS did not record the figures that we needed. While 2011-2018 had data on percentage of *family households* that were led by single females, 2019 data only had data on total households led by single females and total households led by single females with children. We chose the latter, and as the reader can tell, the 2019 percentages are much lower than the 2011-2018. We decided that incorporating such pre-treatment data would skew the synthetic control pretreatment averages too much and decided to cut that data.

[3] We tracked percentage of the population who received SNAP benefits, but we did not use that data for any purpose. That variable was also not used by the Wells and Weishelt study.

Alexandria, VA	4.2	18.8	43.3	15	10817.200	51	1
Alexandria, VA	4.2	18.8	43.3	15	10768.333	52	1
Alexandria, VA	4.2	18.8	43.3	15	10768.333	53	1
Ann Arbor, MI	6.3	12.0	46.4	28	4066.000	108	2
Ann Arbor, MI	6.8	13.5	45.5	28	4107.429	109	2
Ann Arbor, MI	7.6	13.2	45.7	28	4171.393	110	2
Ann Arbor, MI	7.6	14.3	45.7	28	4206.000	111	2
Ann Arbor, MI	7.4	13.6	44.8	28	4240.357	112	2
Ann Arbor, MI	6.4	13.3	45.0	28	4203.143	113	2
Ann Arbor, MI	5.9	10.8	45.9	28	4354.643	114	2
Ann Arbor, MI	5.2	11.3	44.8	28	4377.536	115	2
Ann Arbor, MI	4.9	9.8	45.2	28	4389.036	116	2

We decided to split our analysis into violent and property crime analysis for similar reasons as Wells and Weishelt did in their study. There is no evidence that violent and property crime trends are parallel, and ordinarily, property crime numbers would constitute 90% of the total crime rate. Additionally, because the UCR primarily reports property and violent crime numbers, property and violent crime numbers were already standardized before we began analyzing the data.

Methodology A: Approximating the Ideal Synthetic Control Methodology

In our first methodology, we analyzed 3 treated jurisdictions (Denver, Colorado Springs, and Aurora) and included 88 control jurisdictions. The treated jurisdictions were chosen on the basis that they were the 3 largest cities within Colorado. We excluded cities in Connecticut, as they passed their own version of qualified immunity reform. We did not need to further exclude New Mexico and New York City, since such jurisdictions were missing data and did not appear in our final dataset.

Using the Synth package, we created synthetic controls of each of the three treated jurisdictions for both violent and property crime rates. We tested a series of synthetic controls to determine the helpfulness of particular predictor variables in the analysis, but we ended up keeping all predictor variables that we mentioned earlier to preserve methodological standardization.

When creating synthetic controls, we included all pre-treatment time periods but optimized over 2012 to 2019, allowing the "Synth" function to automatically calculate the pre-treatment mean squared prediction error ("MSPE") over those eight years. We specified the pre-treatment time period to be 2011 to 2020. Although 2020 was the year that the qualified immunity law was passed in Colorado, the function we used to calculate MSPE ratios was the "generate.placebos" function from the SCtools package, which included the final pre-treatment year and the post-treatment years in calculating the post-treatment MSPE. Thus, although the pre-treatment time period was specified to be 2011 to 2020, for functional purposes, 2012 to 2019 were the years relevant to the pre-treatment MSPE calculation, and 2020-2021 were the years relevant to the post-treatment MSPE calculation. Additionally, we specified for the function to employ every available optimization method and choose the best-performing method.[4] We ended by creating

[4] Methodologically, it may have been stronger to stick to one optimization method to standardize calculations and reduce computing times. However, when running the Synth function, we sometimes received errors ("Error in svd(c): Infinite or missing values in 'x'") which resulted from optimization methods sometimes producing matrices with 0s. To stop producing these errors,

six different synthetic controls, two for each Colorado jurisdiction, and within each jurisdiction, one for violent crime rates and one for property crime rates.

To determine the significance of our findings, we calculated the MSPE ratio for each of the synthetic controls.[5] In other words, we averaged the squared amount that the synthetic violent or property crime rates differed from the observed violent or property crime rates over the optimized pretreatment time period, given by the Synth function as the loss.v value. We then averaged the squared amount that the synthetic violent or property crime rates differed from the observed violent or property crime rates over the post-treatment time period. To control for jurisdictions where the synthetic model was not a great fit, we divided the post-treatment MSPE by the pre-treatment MSPE, creating an MSPE ratio. Theoretically, if the intervention had a significant effect on the property or violent crime rates in the treated jurisdictions, we should see significant increases in crime rates in 2020 and 2021 exceeding those of the synthetic control, and thus, the MSPE ratio of those jurisdictions should be high. However, because there is no objective metric for what a "high enough" MSPE ratio is, we created placebo synthetic controls for every control jurisdiction in the dataset and calculated MSPE ratios for each placebo synthetic control. If the MSPE ratio of the treated jurisdiction was greater than 95% of placebo MSPE ratios, we concluded that the MSPE ratio of the treated jurisdiction was high enough to be statistically significant. The interpretations of such a conclusion are further discussed in the "Discussion" section.

We test for two primary hypotheses:

1. The passage of SB-217 coincided with statistically significant gaps in violent crime rates between all three treated jurisdictions and their controls when standardized for pre-treatment fit. Statistical significance is quantified using placebo MSPE ratios. Further, the treated jurisdictions' violent crime rates are *greater than* their synthetic controls.

2. The passage of SB-217 coincided with statistically significant gaps in property crime rates between all three treated jurisdictions and their controls when standardized for pre-treatment fit. Statistical significance is quantified using placebo MSPE ratios. Further, the treated jurisdictions' property crime rates are *greater than* their synthetic controls.

If the evidence proves either hypothesis true, the data would provide some evidence (though not conclusive) for a plausible causal chain between SB-217 and higher crime rates. Because we are testing if a statewide causal factor (SB-217) explained the increase, the hypotheses are only proven true if *all three* treated jurisdictions have significant MSPE ratios. Denver and Aurora, alone, do not provide enough evidence because they belong to the same metropolitan statistical

---

we were forced to run all optimization methods, even if such a method increased computing times significantly when generating placebos.

[5] "MSPE" refers to mean squared prediction error, a measure of how well a model matches the observed outcome variable. A higher MSPE generally indicates more "error," meaning the model's predictions significantly deviate from reality. Generally, in a synthetic control methodology, we wish to minimize pre-treatment MSPE (or the MSPE before the date of the policy intervention) to obtain a better fit. However, high post-treatment MSPE may indicate that the policy intervention had an observable effect on the outcome variable in the jurisdiction, as the jurisdiction's true values differed substantially from those expected by the control. We use MSPE ratio, or the post-treatment MSPE divided by pre-treatment MSPE, to express how much the observed values differ from what we expect based on the model, controlled for how well the model fit prior to treatment.

area and are expected to have similar trends. If Colorado Springs does not experience a significant increase and Denver and Aurora experience a significant increase, the data would only provide evidence for a local causal factor driving up crime, not a statewide causal variable.

The synthetic control methodology that we employed has several limitations. First, as noted in the "Sampling Methodology" section, we extrapolated predictor values from 2019 to 2020 and 2021. This biased the averages used when constructing the synthetic control. Since 2020 was included in the pre-treatment time section, we functionally doubled the role of 2019 in calculating predictor averages for the treated jurisdiction for the synthetic jurisdiction to emulate. We do not believe this should, alone, invalidate our analysis. Since the 2020 predictors data is only used in calculating an overall average of the predictors that the synthetic jurisdiction should approximate - and not to serve as predictors that should be held constant from year to year to isolate the effects of the intervention - the extrapolated 2020 data would only cause the synthetic control methodology to create weighted averages that matched treated jurisdictions' 2019 data above other earlier years. For example, if researchers attempted to control the 2020 and 2021 MSPE for the predictor variables using the 2019 data, such an effort would clearly be invalid, as 2019 unemployment and child poverty rates cannot be used to adjust for 2020 and 2021 data. However, because we do not calculate MSPEs differently based on predictor values, we do not suffer from such limitations. The methodology merely averages the predictor values of the treated jurisdiction over the pre-treatment time period for the synthetic control to match but does not attempt to hold such predictors constant from year to year or control for yearly shifts in those predictors. Thus, any skew created by such a flaw is minimal.

Second, the time of the treatment is not effectively accounted for by the synthetic control methodology. The passage of the police accountability bill in Colorado occurred in the middle of 2020; however, we do not have quarterly data by which we could isolate the two quarters of 2020 prior to treatment from the two quarters post-treatment. Instead, we simply sort 2020 and 2021 as broadly falling under the post-treatment time frame, operating on the assumption that if the police accountability legislation affected violent and property crime rates in the Colorado jurisdictions, the increase in violent and property crime rates for the whole of 2020 would be greater than those of treated jurisdictions. Unfortunately, such an assumption is not necessarily true, as 2020 introduced a series of different factors, ranging from the COVID-19 pandemic to the George Floyd protests, each of which influenced jurisdictions' crime rates in unknown ways. As a result, we are hesitant to derive a causal conclusion from any of our analysis. We attempt to solve this problem in Methodology B at the cost of other significant methodological limitations.

Third, in an ideal synthetic control, we would have a wealth of years both before and after the treatment to evaluate. Unfortunately, due to the recency of the legislation and the inability to divide years into quarterly data, we only had a total of 2 post-treatment time periods to evaluate. This may limit our insights, as a single year of increased property or violent crime rates in one of the treated jurisdictions would skew the mean post-treatment MSPE substantially, even if such a year occurred merely from chance. Placebo testing should diminish the influence of chance in the analysis, but having more post-treatment time periods to calculate the MSPE would allow the analysis to be more reliable.

Methodology B: Synthetic Control as Comparative Case Study Selection

Our second methodology was employed before the release of 2020 and 2021 data by the UCR and was meant to serve as a workaround to normal synthetic analysis. As a result, the second methodology suffers from severe limitations, many of which could invalidate the analysis entirely. We include the results from Methodology B in the appendices in case they are found to be useful in their treatment date precision and high jurisdiction inclusion that Methodology A lacks.

In our second methodology, we examined violent and property crime rates in Denver only.[6] In selecting possible control jurisdictions, we waived the requirements for 2020 and 2021 crime data, as such data was not relevant for the analysis. We additionally only filtered for jurisdictions greater than 50000 in population, as we only had access to a small number of time periods but an enormous sample of jurisdictions within the donor pool. We removed population density from the analysis and relied on population alone to serve as the "population" level statistic. This led to many nonsimilar jurisdictions being included in the analysis, significantly increasing the potential for bias. In total, we had roughly 530 jurisdictions in the donor pool when constructing the synthetic control.

To account for the lack of 2020 and 2021 data and to increase the precision of the treatment dates, we used the synthetic control methodology to identify jurisdictions similar to Denver and to provide weights for some of those jurisdictions. We optimized the synthetic controls for 2016 to 2019 to obtain a synthetic control that could follow the most recent trends in Denver. We then identified the top five jurisdictions with the highest weights and reran the synthetic control model with only those jurisdictions to recalculate the weights, relying on the premise that the synthetic of the top four or five jurisdictions that comprise a majority weight in the full synthetic control would be similar enough to the treated jurisdictions to analyze. With those identified control jurisdictions, we submitted requests for incident-level crime data to 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.

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, burglary, 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 2019 to June 2020 (before qualified immunity reform) and June 2020 to June 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 property crimes proportionate to each jurisdiction's respective crime numbers. Finally, we created a bootstrapped null distribution assuming no true difference between the

[6] We also attempted to analyze Colorado Springs using this methodology, but we had trouble requisitioning the needed data in a useable form from police agencies.

daily increases of the synthetic jurisdiction compared to Denver and calculated a p-value based on the probability of observing the real difference or greater between Denver and the synthetic control difference based on the null distribution.

This methodology had a few critical limitations. First, because 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 extrapolate beyond the capabilities of the synthetic control. Second, because of several denied requests (in particular, Ann Arbor and Clarksville Police Departments), we were forced to rely partially on convenience sampling in order to successfully carry out the study. 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 is statistically invalid because it assumes some contiguous relationship between corresponding days on different years, where increases from one day to the corresponding day on the next year would have meaning. However, such an assumption is clearly incorrect, as crime numbers on June 14, 2020 are wholly unrelated to crime numbers on June 14, 2021. As a result, this method 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.

Fourth, because we had to standardize the daily crime numbers by dividing crime numbers from some relative figure for each jurisdiction (in this case, the total number of offenses in the 2019-20 time period), smaller jurisdictions disproportionately influenced the variance of the synthetic control, since daily fluctuations of 1-2 offenses were much greater when standardized compared to larger jurisdictions. Fifth, the methodology misuses the synthetic control methodology to identify 4-5 jurisdictions that comprise the majority weight of the main jurisdiction, but the synthetic control methodology is only intended to weight jurisdictions in a manner that creates an average jurisdiction matching the treated one, not to identify jurisdictions that are most similar to the treated jurisdiction. As a result, the jurisdictions we chose based on the synthetic control were often dramatically different from the treated jurisdiction (such as Champaign, IL and Fort Smith, AR, both of which were incredibly small jurisdictions). Finally, because we included both excessively small and excessively large jurisdictions, we did not filter the dataset beforehand to only include jurisdictions that were somewhat similar to Denver, skewing the synthetic averages towards the extremes.

The results of this methodology, in their entirety, are described in Appendix 2.

Exploratory Data Analysis

In this section, we lay out crime numbers from our data and compare them with crime numbers generated by synthetic controls derived from Methodology B to give readers an idea of what conclusions we expected prior to running the analysis. We did not include Aurora graphs in the Exploratory Data Analysis, since we did not have access to incident-level Aurora data. We expect, however, that Aurora's crime numbers parallel Denver's.

As we noted before, if the statewide police accountability law led to systematic increases in crime rates, we would expect to see roughly parallel increases across both Denver and Colorado

Springs. One city experiencing upticks in crime that the other city does not experience only provides evidence of a local causal mechanism, not a statewide causal factor.

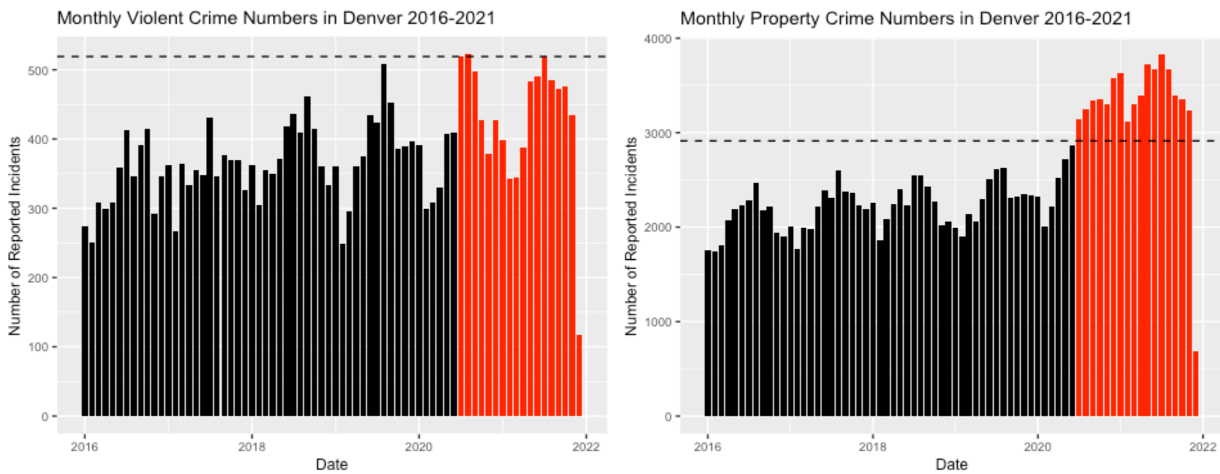


Fig. 1: Monthly Violent and Property Crime Numbers in Denver 2016-2021: Each bar represents the number of reported incidents in a single month. Red bars represent the months following the passage of the police accountability legislation on June 19, 2020. Data from June 20-30 is included in the month immediately preceding the red bars (June 2020). The black dotted line represents 10 reported violent incidents or 50 reported property incidents above the previous maximum number of offenses in a single month in the four years prior to legislation.

After the passage of the police accountability law on June 19, 2020, Denver experienced some increase in violent crimes. Both July and August 2020 had more violent crimes in a single month than the previous four years' record for violent crimes in a single month. Denver's violent crimes then decreased over the fall and winter before increasing again the following summer, reaching similar crime numbers as the previous summer. We could interpret Denver's violent crime increase as part of Denver's steady yearly increases in violent crime since 2016.

On the other hand, Denver's property crime incidents increased far more dramatically than its violent crimes did. In **every month** following the passage of the police accountability legislation, Denver experienced more property crimes than the city had in any single month in the previous 4-5 years. Denver's property crimes also did not decrease to normal levels as Denver's violent crimes did. Importantly, however, Denver's property crime increase seems to have begun around March or April 2020, not in June, possibly implying that other factors (such as the COVID-19 pandemic) may have fueled the rise in property crime.

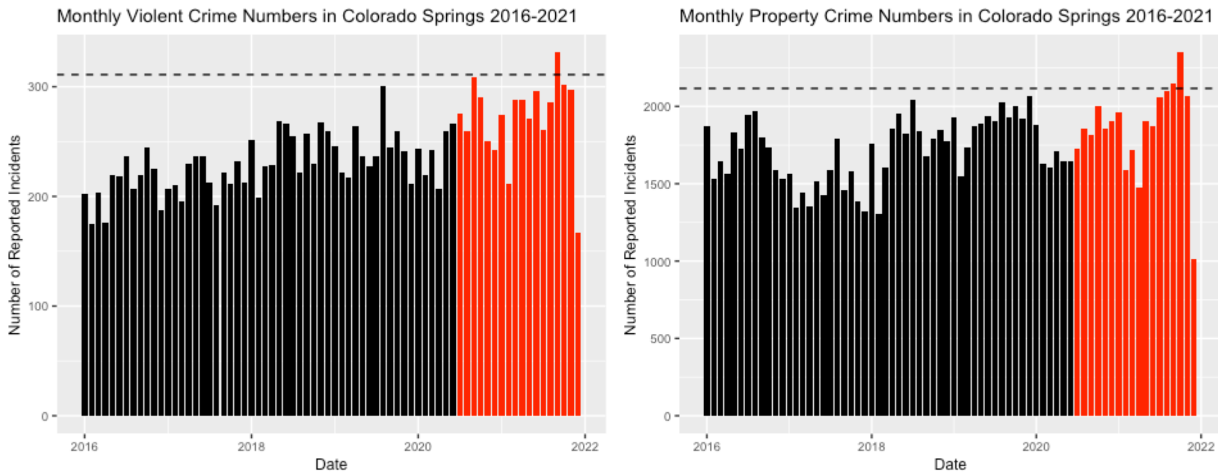


Fig. 2: Monthly Violent and Property Crime Numbers in Colorado Springs 2016-2021

Similarly to Denver, Colorado Springs also experienced some increase in violent crime following the passage of the police accountability legislation. In the summer of 2020, Colorado Springs experienced moderately high violent crime, roughly matching the heights of the previous summer. Additionally, in the summer of 2021, Colorado Springs' violent crime numbers increased significantly, with one month far exceeding the single-month record for number of violent crimes from the past 4 years.

Colorado Springs also appears to have experienced some rise in property crime in the summer of 2021, although the increase is not nearly as pronounced as the increase that Denver experienced. The summer of 2020 did not appear to have unusually high property crime numbers. The graph does not present clear evidence that Colorado Springs' property crimes substantially increased following the passage of the police accountability legislation. As we noted earlier, if the rise in property crime in both jurisdictions was caused by the police accountability legislation, we would expect to see roughly parallel trends in both jurisdictions instead of the outcome lag and much smaller magnitude increase in Colorado Springs.

To give a control standard for reference, we included the following graphs from Methodology B comparing the monthly weighted averages of 4-5 jurisdictions with Denver's monthly violent crime and property crime rates. This is not data used in our main analysis.

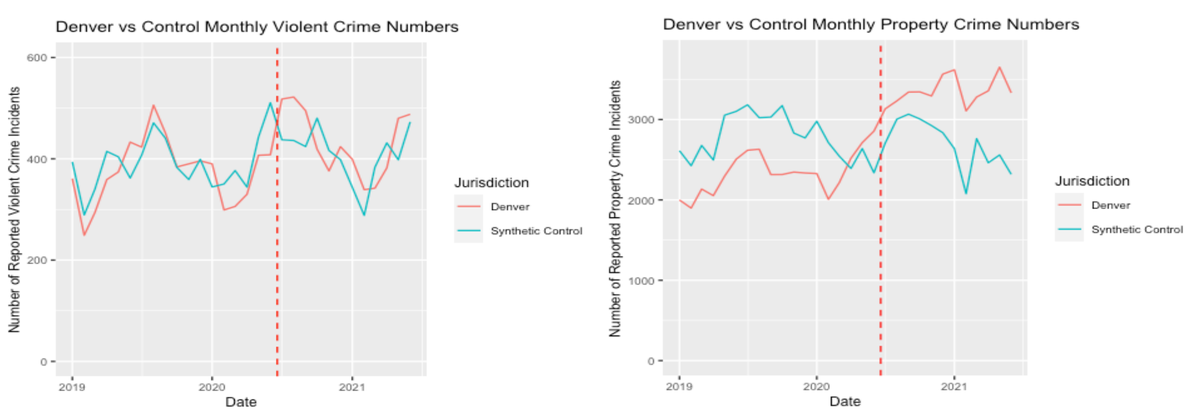


Fig. 3: Monthly Violent and Property Crime Numbers in Denver Compared to Control 2019-2021: The red dashed line represents the passage of the police accountability legislation in June 19, 2020. Although the control

continues to mostly track Denver violent crimes even after the passage of the police accountability law, Denver property crimes far surpass the control after April of 2020.

The violent crime control closely tracks Denver's monthly numbers both before and after the passage of the police accountability law. Although Denver experienced some increases in the summer immediately following the passage of the police accountability bill that were not fully matched by the synthetic control, Denver's numbers soon fell comfortably into the control model's range.

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 crimes also remained high even after the summer, maintaining its much higher position compared to the synthetic control even as late as June of 2021.

In conclusion, the above graphs imply the following possible results. First, Colorado jurisdictions experienced *some* increase in violent crime rates following the passage of police accountability legislation, but those increases may not be large enough in magnitude for chance to be ruled out as a plausible explanation. Second, Denver experienced an extreme increase in property crime rates in the summer of 2020 that never decreased to normal levels, implying a high likelihood that Denver's property crime increase is *sustained* and *due to systematic factors other than chance*. On the other hand, while Colorado Springs experienced some increase in property crime rates, its significance is questionable due to its much lower magnitude.

Testing and Results

In this section, we discuss the synthetic control diagnostics and MSPE test results for each of the six synthetic controls.

Synthetic Controls for Violent Crime

Denver

We began with Denver, the largest jurisdiction in Colorado. Below, we included the graph comparing the violent crime rates of the synthetic jurisdiction and Denver itself. We also included a table comparing observed and synthetic predictor values to evaluate model fit.

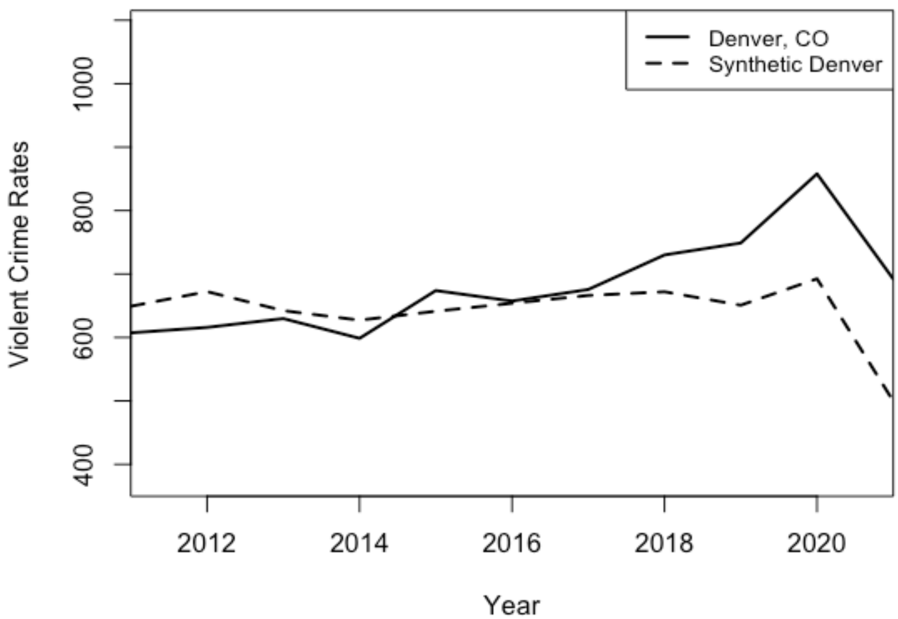


Fig. 4: Denver vs Synthetic Control Violent Crime Rates 2011-2021: The model fit exceptionally well from 2011-2017 before some declining fit in 2018 and 2019. 2020 was the recorded year of treatment.

Table 2: Observed vs Synthetic Denver Predictor Values

	Treated	Synthetic	Sample Mean
Population	682917.900	682732.925	312485.499
Population Density	4463.516	4463.030	3832.261
Median Income (USD)	56967.300	56952.416	55733.756
HS Education or Above %	86.310	86.312	86.464
Residential Stability %	78.240	78.244	80.522
Over 18 %	79.270	77.703	76.249
White %	53.200	53.181	53.491
Self-Employed Rate	5.640	5.640	5.282
Unemployment Rate	6.120	6.135	7.560
Owner-Occupied Housing %	50.030	50.023	53.687
Child Poverty Rate	24.340	24.338	23.744

The synthetic control is relatively strong. The synthetic control matches observed Denver's predictor values exceptionally well, and the pre-treatment MSPE is relatively low at approximately 2288.114 (as computed by the Synth package).

Based on the graph, we can conclude that Denver's violent crime rates were greater than control jurisdictions in the post-treatment period. However, it is difficult to tell whether Denver's increase in violent crime rates in 2020 is due to systematic causal factors in Denver (like the treatment) or simply due to declining fit in the more recent years. While synthetic Denver roughly follows the trends of Denver up until about 2017, synthetic Denver begins diverging from observed Denver as early as 2018. The pre-treatment gap between the predicted and observed values only increases in 2019 before Denver's large increase in 2020.

We include the following table to be transparent about how the synthetic control ended up assigning the largest weights.

Table 3: Weights of the Top 20 Highly Weighted Jurisdictions

Weights	Unit Names
0.279	Oklahoma City, Oklahoma
0.158	Houston, Texas
0.114	Alexandria, Virginia
0.067	Cambridge, Massachusetts
0.061	Springfield, Missouri
0.040	Madison, Wisconsin
0.035	Austin, Texas
0.014	Columbia, Missouri
0.014	Fargo, North Dakota
0.008	Dallas, Texas
0.008	Laredo, Texas
0.008	Waco, Texas
0.008	Wichita Falls, Texas
0.007	Ann Arbor, Michigan
0.007	College Station, Texas
0.006	Lexington, Kentucky
0.005	Evansville, Indiana
0.005	Manchester, NH
0.005	Odessa, Texas
0.005	Salt Lake City, Utah

In the weights, we can see that the majority of synthetic Denver is comprised of Oklahoma City, Houston, and Alexandria, although several cities possess nonzero weights. In this way, our synthetic control is distinct from the synthetic control used within Abadie et. al.'s research on the Basque region, as their synthetic control only weighted 2 regions and assigned zero weights to the rest of the regions.

Results of Placebo Testing: After calculating 88 different placebo synthetic controls, we found that Denver's MSPE ratio was nonsignificant at the 1%, 5%, or 10% level. When excluding jurisdictions with pre-treatment MSPEs more than 5x greater than the pre-treatment MSPE of Denver, more than 11% of the placebo synthetic controls had MSPE ratios greater than that of Denver. However, Denver's data is still relatively extreme; if we instead chose to run a one-sided test or excluded outliers from the placebos, it is very plausible that Denver's violent crime rates would be significant. Nonetheless, based on our assigned thresholds, the data does not provide sufficient evidence to rule out chance as an explanation for the differences in violent crime rates between Denver and the synthetic control in 2020 and 2021.

Colorado Springs

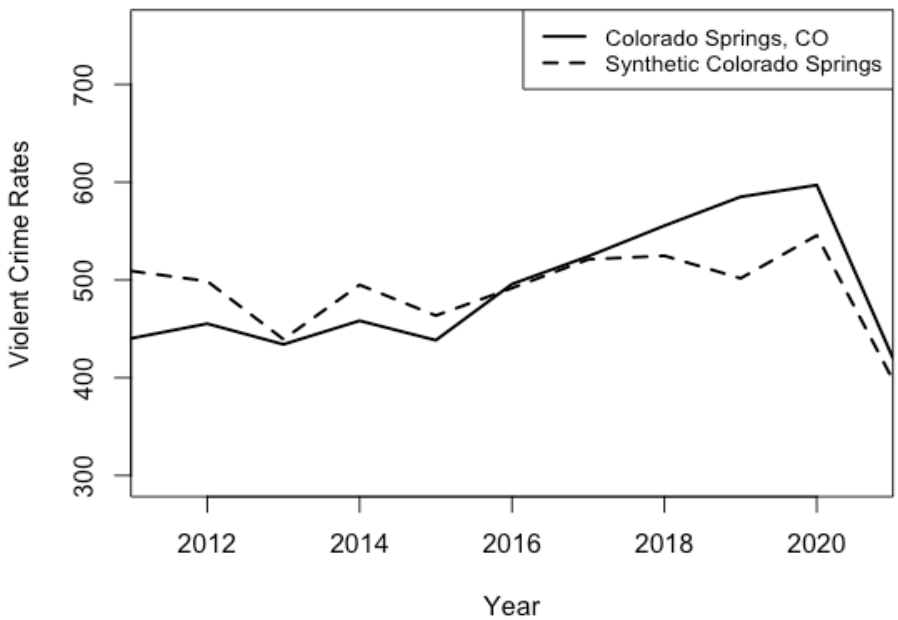


Fig. 5: Colorado Springs vs Synthetic Control Violent Crime Rates 2011-2021: The model fits relatively well until 2018, where the gap between violent crime rates increases significantly just prior to treatment.

Table 4: Observed vs Synthetic Colorado Springs Predictor Values

	Treated	Synthetic	Sample Mean
Population	456236.000	450584.230	312485.499
Population Density	2339.672	2367.789	3832.261
Median Income (USD)	59.040	55.666	53.687
HS Education or Above %	93.160	90.760	86.464
Residential Stability %	76.380	76.489	80.522
Over 18 %	75.880	75.874	76.249
White %	69.540	58.711	53.491
Self-Employed Rate	5.550	5.404	5.282
Unemployment Rate	7.790	7.685	7.560
Owner-Occupied Housing %	57594.800	57654.295	55733.756
Child Poverty Rate	17.520	19.241	23.744

The synthetic control for Colorado Springs is very strong. The pre-treatment MSPE value is approximately 1477.022, lower than that of Denver. Per Table 4, Colorado Springs' predictor values match relatively well with those of the synthetic jurisdiction, with the exceptions of racial homogeneity (about 10% off), child poverty (about 1.8% off), and owner-occupied housing (about 3% off). Nevertheless, given that the synthetic control tracks Colorado Springs relatively thoroughly, we find it appropriate to proceed with the given model.

Based on the graph, Colorado Springs did have slightly higher violent crime rate values than the synthetic control in both 2020 and 2021. However, we are relatively certain that such a gap is explainable by declining fit in the later years. The largest violent crime gap is in 2019, where Colorado Springs experienced an increase in violent crime rate while the synthetic control experienced a decrease. The gap decreases in both 2020 and 2021, implying that the only reason violent crime rates are "higher than expected" is because they were already higher pre-treatment.

Table 5: Weights of the Top 20 Highly Weighted Jurisdictions

	Weights	Unit Names
17	0.403	Clarksville, Tennessee
70	0.183	Peoria, Arizona
2	0.170	Ann Arbor, Michigan
39	0.138	Houston, Texas
13	0.047	Carlsbad, California
34	0.015	Frisco, Texas
8	0.004	Boise, Idaho
38	0.004	Henderson, Nevada
25	0.002	Detroit, Michigan
49	0.002	Las Vegas, Nevada
60	0.002	Mesa, Arizona
77	0.002	San Antonio, Texas
21	0.001	Corpus Christi, Texas
23	0.001	Dayton, Ohio
33	0.001	Fort Wayne, Indiana
35	0.001	Garland, Texas
47	0.001	Lansing, Michigan
50	0.001	League City, Texas
51	0.001	Lee's Summit, Missouri
57	0.001	McAllen, Texas

Per Table 5, roughly 90% of the synthetic control weight is centered around 4 jurisdictions: Clarksville, Peoria, Ann Arbor, and Houston. The rest of the jurisdictions have weights just above 0, similar to the weights we expected from Abadie et. al.'s analysis.

Results of Placebo Testing: Using the same placebos generated for Denver, we found that Colorado Springs' MSPE ratio of 1.066 was smaller than 70% of placebo jurisdictions after removing placebo jurisdictions with pre-treatment MSPEs five times greater than that of Colorado Springs. The data does not provide sufficient evidence to indicate that Colorado Springs' violent crime rates in 2020 and 2021 were significantly different from those of control jurisdictions.

Aurora

We expect to see roughly the same crime trends in both Aurora and Denver. Below, we depict a plot comparing the observed/synthetic violent crime rates as well as a plot depicting the gaps in greater detail to visualize the weak model fit more easily.

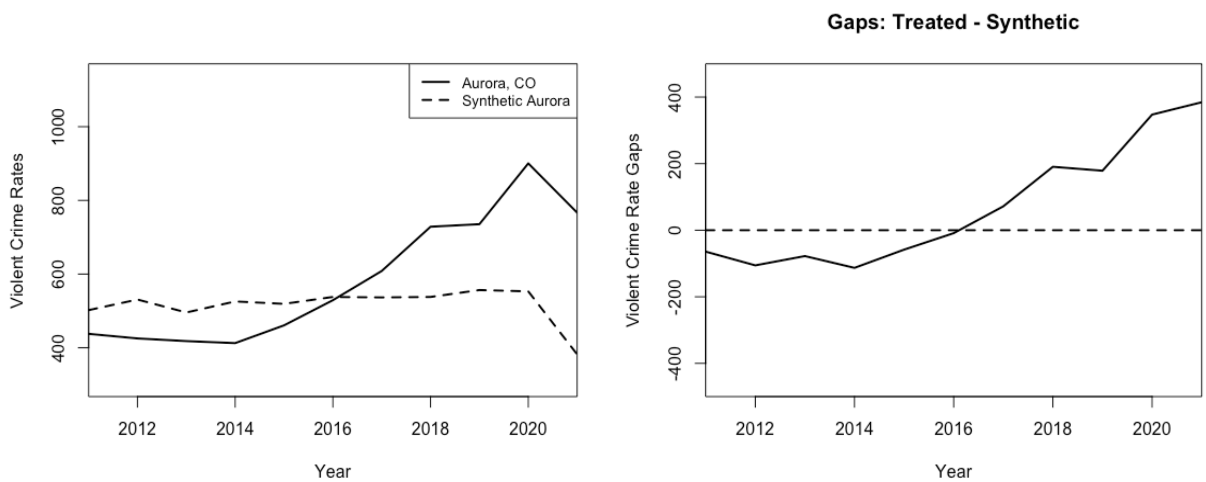


Fig. 6: Aurora vs Synthetic Control Violent Crime Rates 2011-2021: The path plot on the left follows the violent crime rate numbers in the observed and synthetic jurisdictions. The gaps plot subtracts the synthetic (expected) violent crime rates from the observed violent crime rates to show the numerical gaps between the jurisdictions over time. Unlike the first two synthetic controls, the model fit is extremely weak.

Table 6: Observed vs Synthetic Aurora Predictor Values

	Treated	Synthetic	Sample Mean
Population	359600.00	357346.825	312485.499
Population Density	2320.00	2348.947	3832.261
Median Income (USD)	58.64	58.015	53.687
HS Education or Above %	86.62	86.663	86.464
Residential Stability %	79.03	79.127	80.522
Over 18 %	73.61	73.626	76.249
White %	46.09	49.700	53.491
Self-Employed Rate	5.00	5.034	5.282
Unemployment Rate	7.56	7.551	7.560

Owner-Occupied Housing %	56417.60	56446.994	55733.756
Child Poverty Rate	20.51	20.886	23.744

Even though Aurora's synthetic control matches its predictor values extremely well, the graph shows that the model is a weak fit for the data. While Aurora experiences a large increase in violent crime in 2016 and 2017, the synthetic control experiences no such increase. Aurora's pre-treatment MSPE is also high with a value of approximately 13362.431, almost 10 times that of Colorado Springs. We found it unlikely that insights derived from this synthetic control would be helpful, but we ran the significance test regardless.

As in the other two jurisdictions, Aurora's violent crime rate post-treatment is greater than the synthetic control. However, as the gaps plot demonstrates, the gap between Aurora and the synthetic control had been steadily increasing for some time before increasing dramatically post-treatment. It is plausible that the increase resulted from the police accountability law, but more likely, the increased gap in 2020 and 2021 was simply a symptom of the already weak model fit and preexisting violent crime trends in Aurora.

Table 7: Weights of the Top 20 Highly Weighted Jurisdictions

Weights	Unit Names
0.254	Clarksville, Tennessee
0.163	Chesapeake, Virginia
0.138	San Antonio, Texas
0.092	Odessa, Texas
0.082	Pasadena, Texas
0.061	Round Rock, Texas
0.035	College Station, Texas
0.024	Kenosha, Wisconsin
0.011	Frisco, Texas
0.008	Olathe, Kansas
0.006	Columbia, Missouri
0.005	Grand Prairie, Texas
0.005	Waco, Texas
0.004	Fort Wayne, Indiana
0.004	Green Bay, Wisconsin
0.003	El Paso, Texas
0.003	Fargo, North Dakota
0.003	Irving, Texas
0.003	Las Vegas, Nevada
0.003	League City, Texas

Results of Placebo Testing: We found that Aurora's MSPE ratio is not significant at the 1%, 5%, or 10% level. When excluding placebos with pre-treatment MSPEs over five times greater than that of Aurora, about 11.9% of the placebo synthetic controls have MSPE ratios greater than that of Aurora. It is plausible that if we distinguished placebo MSPE ratios with higher than expected violent crime rates from placebo MSPE ratios with lower than expected violent crime rates (functionally turning the test into a one-sided test), Aurora's increase may become significant. However, given that Aurora's synthetic control is already so weak, we do not feel it would be valuable to conduct such an analysis.

Placebos for Violent Crimes

As we noted earlier, we generated 88 different placebo synthetic controls with the same settings as the original synthetic control and compiled all the MSPE ratios from each placebo synthetic control into a single dataset for significance testing. We would like to take a moment to comment on these placebos.

Below, we visualized these placebos' MSPE ratios and noted their summary statistics.

Table 8: Summary Statistics of Violent Crime Placebos (with Outliers)

Mean	Standard Deviation	Median
8.043	25.021	1.641

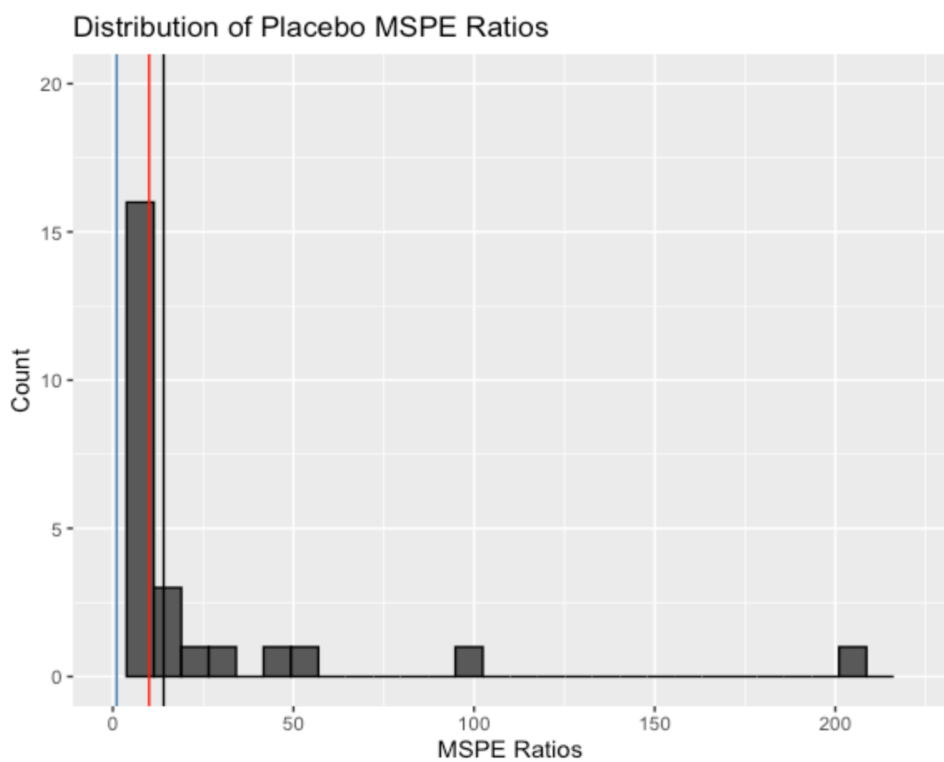


Fig. 7: Placebo MSPE Ratios for Violent Crime Rates: Here, we displayed a histogram of all MSPE ratios calculated by the "generate.placebos" command. The black line represents Denver's MSPE ratio, the red line represents Aurora's MSPE ratio, and the blue line represents Colorado Springs' MSPE ratio. We calculate p-values by dividing the number of "more extreme" MSPE ratios (to the right of the lines) by the total number of MSPE ratios.

As the histogram displays, most placebo MSPE ratios are centered at 0-10 with the exception of two placebo MSPE ratios above 80 and a series of other outliers in the 20-50 range. Those extreme outlier MSPE ratios represent Evansville and Manchester and likely occurred from an exceptionally good fit with the data in pretreatment years with some declining fit in 2020. The outliers are further discussed in Appendix 2. The summary statistics in Table 8 further demonstrate how much the outliers differ from the rest of the dataset; while the median is centered on an MSPE ratio of around 1, the standard deviation is 25 and the mean is 8.

Re-visualizing without outliers and reducing the binwidth to further detail the smaller MSPE ratios in the spectrum, we arrive at the second graph and table below.

Table 9: Summary Statistics of Violent Crime Placebos (without Outliers)

Mean	Standard Deviation	Median
4.73	8.649	1.609

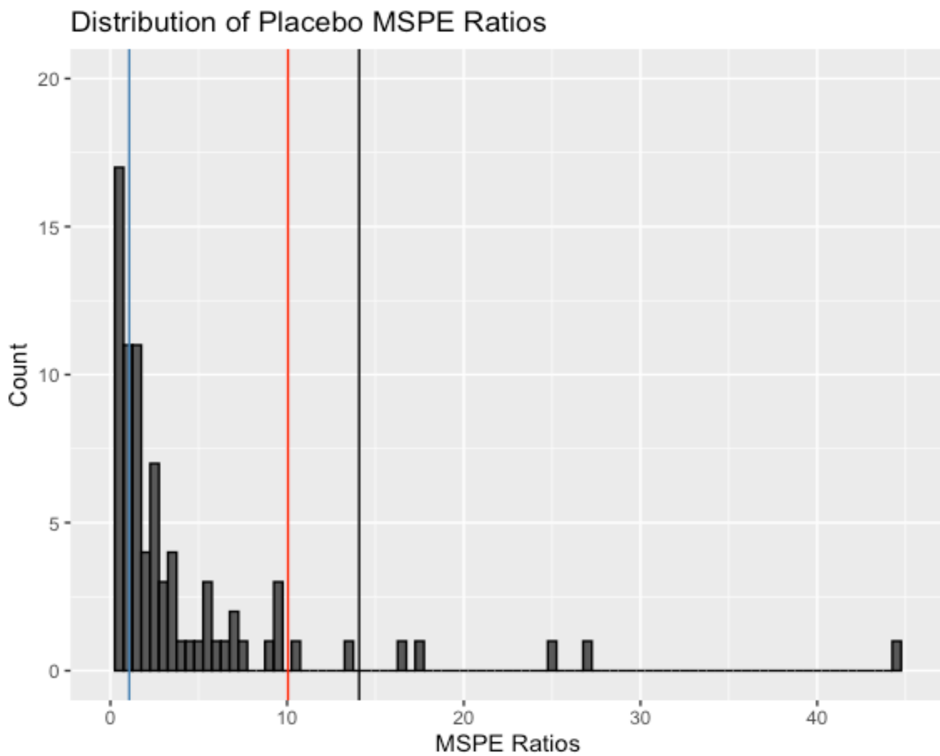


Fig. 8: Placebo MSPE Ratios for Violent Crime Rates without Outliers: The black line represents Denver's MSPE ratio, the red line represents Aurora's MSPE ratio, and the blue line represents Colorado Springs' MSPE ratio. We calculate p-values by dividing the number of "more extreme" MSPE ratios (to the right of the lines) by the total number of MSPE ratios.

As we can see, Colorado Springs is squarely within the center of the distribution. On the other hand, Aurora and Denver's MSPE ratios are larger than most of the MSPE ratios within the dataset, but the ratios are still smaller than enough placebos to not constitute statistically significant evidence.

We display these placebos to give the reader an idea of what synthetic controls we ended up creating, any outliers or flaws within the synthetic controls, as well as where the treated jurisdictions lie on the distribution.

Overall Results for Violent Crime

We reject the first hypothesis. After generating 88 placebo MSPE ratios, we found that the MSPE ratios of Denver, Colorado Springs, and Aurora were not large enough to constitute statistically significant evidence that violent crime rates in those areas were significantly different from violent crime rates in control jurisdictions post-treatment. The data does not provide sufficient evidence that SB-217 coincided with statistically significant increases in violent crime rates compared to control jurisdictions.

Property Crime

Our property crime results varied significantly from our violent crime results in terms of the significance of our findings. Just like in the case of violent crime, we created three synthetic controls, one for each of the three treated jurisdictions. Model fit varied significantly based on the treated jurisdiction.

Denver

Because of weaker fit, we included both the gaps plot and the path plot for the data.

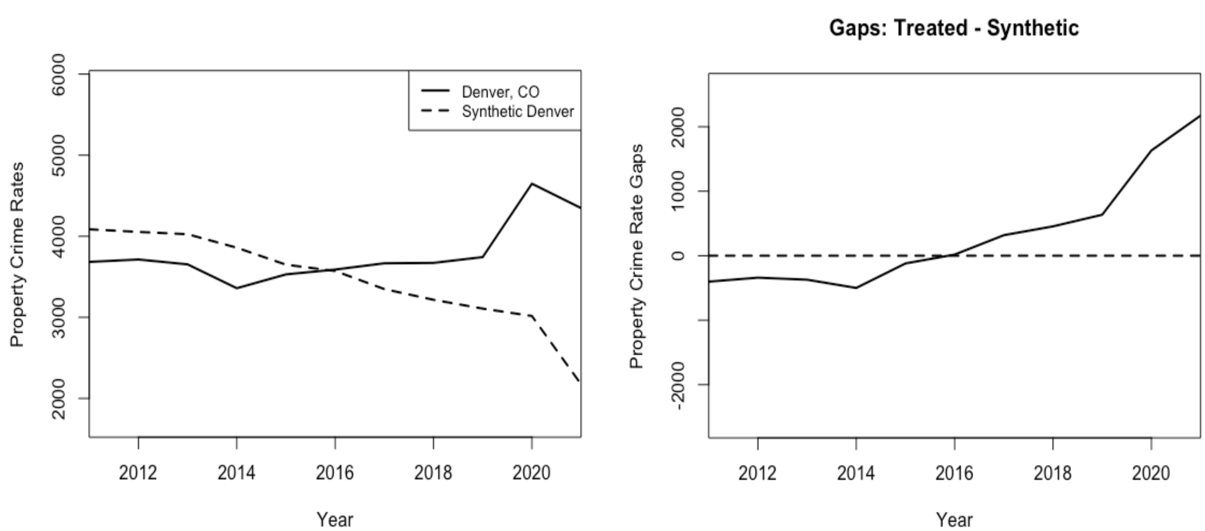


Fig. 9: Denver vs Synthetic Control Property Crime Rates 2011-2021: The path plot on the left follows the property crime rate numbers in the observed and synthetic jurisdictions. The gaps plot subtracts the synthetic (expected) violent crime rates from the observed violent crime rates to show the numerical gaps between the jurisdictions over time.

Table 10: Observed vs Synthetic Denver Predictor Values

	Treated	Synthetic	Sample Mean
Population	682917.900	682770.807	312485.499
Population Density	4463.516	4470.970	3832.261
Median Income (USD)	56967.300	56835.758	55733.756
HS Education or Above %	86.310	86.341	86.464
Residential Stability %	78.240	78.248	80.522

Over 18 %	79.270	79.006	76.249
White %	53.200	53.070	53.491
Self-Employed Rate	5.640	5.638	5.282
Unemployment Rate	6.120	7.620	7.560
Owner-Occupied Housing %	50.030	50.027	53.687
Child Poverty Rate	24.340	24.338	23.744

In terms of predictors, Table 10 displays that the synthetic control does a good job of creating a synthetic jurisdiction with predictors that match observed Denver well. However, Figure 9 displays that the synthetic control is not a strong fit for the Denver data. While the synthetic control's property crime rate steadily decreases, Denver's property crime rate has the opposite trend from 2014. Notably, the gaps between Denver's property crime rates and the synthetic control steadily increase from 2014, although the gaps are not particularly large until the spike in property crimes in 2020.

As the Exploratory Data Analysis led us to expect, Denver's property crime rate is far above the synthetic control's property crime rate post-treatment. There is some declining fit over time, but Figure 9 displays a clear spike in property crime that implies the existence of a systematic causal factor.

Table 11: Weights of Top 20 Highly Weighted Jurisdictions

Weights	Unit Names
0.149	Houston, Texas
0.121	Seattle, Washington
0.103	Ann Arbor, Michigan
0.096	Dayton, Ohio
0.087	San Diego, California
0.064	Sterling Heights, Michigan
0.024	Cambridge, Massachusetts
0.019	Bellevue, Washington
0.016	Dallas, Texas
0.015	Springfield, Missouri
0.013	Manchester, New Hampshire
0.012	Brownsville, Texas
0.011	Escondido, California
0.011	Knoxville, Tennessee
0.009	Alexandria, Virginia
0.009	Austin, Texas
0.008	Lexington, Kentucky
0.008	Waco, Texas

0.008	Wichita Falls, Texas
0.007	Oceanside, California

Results of Placebo Testing: We found the MSPE ratio of Denver to be statistically significant. After excluding jurisdictions with pre-treatment MSPEs greater than five times that of Denver, we found that Denver's MSPE ratio of 23.999 was extremely high, only exceeded by Madison, WI and McAllen, TX. Although not significant at the 1% level, such a finding is significant at the 5% and 10% levels. The data does provide sufficient evidence to indicate that Denver's property crime rate gaps in 2020 and 2021 is significantly greater than those of placebo jurisdictions and makes it unlikely that Denver's heightened property crime rate merely resulted from chance.

Colorado Springs

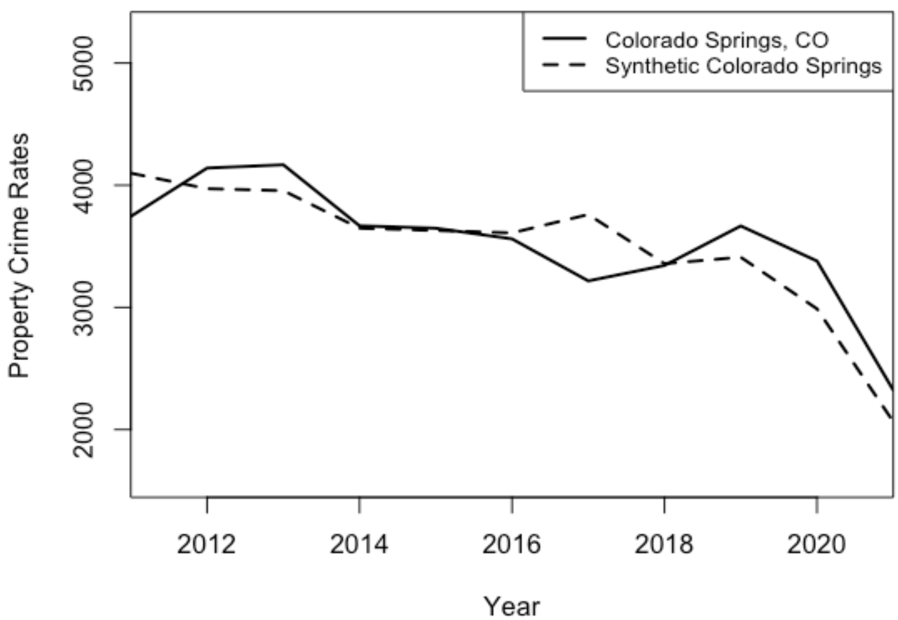


Fig. 10: Colorado Springs vs Synthetic Control Property Crime Rates 2011-2021: The model fits relatively well with a few marginal errors in 2017 and 2019.

Table 12: Observed vs Synthetic Colorado Springs Predictor Values

	Treated	Synthetic	Sample Mean
Population	456236.000	449950.969	312485.499
Population Density	2339.672	2356.828	3832.261
Median Income (USD)	57594.800	55871.890	55733.756
HS Education or Above %	93.160	90.469	86.464
Residential Stability %	76.380	76.568	80.522
Over 18 %	75.880	75.931	76.249

White %	69.540	63.213	53.491
Self-Employed Rate	5.550	5.567	5.282
Unemployment Rate	7.790	7.717	7.560
Owner-Occupied Housing %	59.040	53.574	53.687
Child Poverty Rate	17.520	21.671	23.744

Colorado Springs' synthetic control is a strong fit to the data. With a pretreatment MSPE of only 54693.661 (note that because we are using property crime rates, numbers are expected to be much higher than in the case of violent crime rates), the synthetic control follows Colorado Springs' trend well until about 2019. From 2019-2021, Colorado Springs had a slightly higher property crime rate than its synthetic control. On the level of predictors, synthetic Colorado Springs deviates from Colorado Springs' predictor values in owner-occupied housing, child poverty, and racial homogeneity. This likely limits the extent to which the synthetic control can track Colorado Springs' crime trends effectively, but given the pretreatment fit, a statistical significance analysis would still be meaningful. The weights table in Table 12 are also roughly what we expect, with mostly nonzero weights given and a few jurisdictions comprising the majority of the synthetic control.

Based on Figure 10, Colorado Springs' property crime rate is higher than the synthetic control post-treatment. However, the gap is not very large and likely resulted from the already-present gap in 2019.

Table 12: Weights of Top 20 Highly Weighted Jurisdictions

Weights	Unit Names
0.375	Clarksville, Tennessee
0.172	Springfield, Missouri
0.111	San Diego, California
0.089	Frisco, Texas
0.074	Boise, Idaho
0.069	Lexington, Kentucky
0.064	Houston, Texas
0.028	Ann Arbor, Michigan
0.004	Spokane, Washington
0.003	Las Vegas, Nevada
0.002	Henderson, Nevada
0.001	Columbia, Missouri
0.001	Lee's Summit, Missouri
0.001	Madison, Wisconsin
0.000	Alexandria, Virginia
0.000	Arlington, Texas
0.000	Austin, Texas

0.000	Beaumont, Texas
0.000	Bellevue, Washington
0.000	Brownsville, Texas

Results of Placebo Testing: After removing jurisdictions with pre-treatment MSPEs more than five times greater than that of Colorado Springs, Colorado Springs' MSPE ratio of about 7.7 was not statistically significant at any of the three levels. Roughly 45% of placebo jurisdictions had MSPE ratios greater than the one in Colorado Springs. Thus, the data does not provide sufficient evidence to indicate that Colorado Springs' property crime rate gaps could not have resulted simply from chance.

Aurora

Although the model fit is not weak, we incorporate both the gaps plot and the path plot to display the plausible opposite trends in the data.

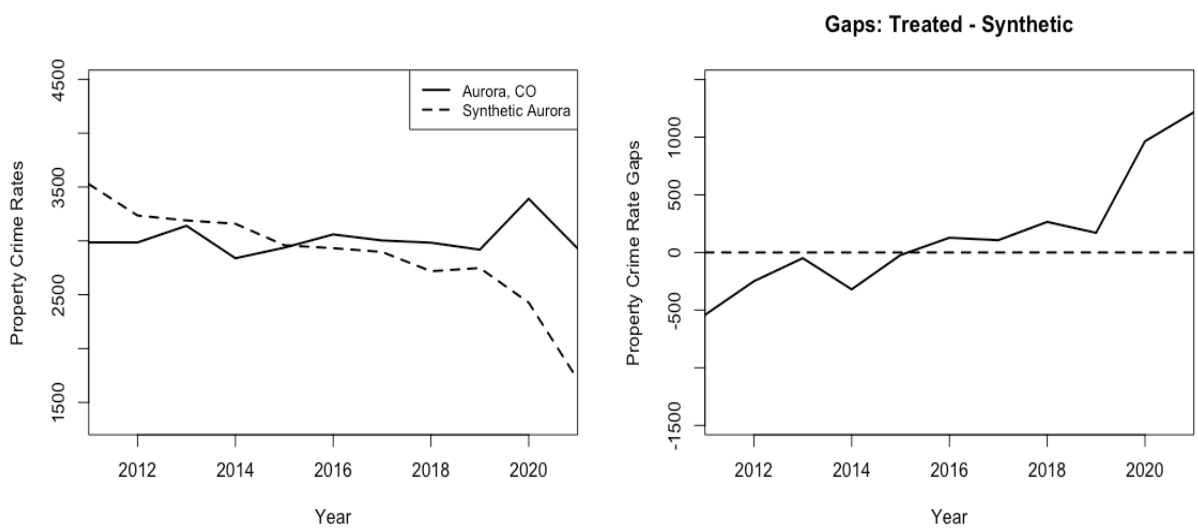


Fig. 11: Aurora vs Synthetic Control Property Crime Rates 2011-2021: The path plot on the left follows the property crime rate numbers in the observed and synthetic jurisdictions. The gaps plot subtracts the synthetic (expected) violent crime rates from the observed violent crime rates to show the numerical gaps between the jurisdictions over time.

Table 13: Observed vs Synthetic Colorado Springs Predictor Values

	Treated	Synthetic	Sample Mean
Population	359600.00	338611.472	312485.499
Population Density	2320.00	2328.506	3832.261
Median Income (USD)	56417.60	56373.063	55733.756
HS Education or Above %	86.62	86.675	86.464
Residential Stability %	79.03	79.047	80.522
Over 18 %	73.61	73.618	76.249

White %	46.09	46.175	53.491
Self-Employed Rate	5.00	5.017	5.282
Unemployment Rate	7.56	7.536	7.560
Owner-Occupied Housing %	58.64	53.943	53.687
Child Poverty Rate	20.51	21.374	23.744

Unlike in the case of the Aurora violent crime synthetic control, the Aurora property crime synthetic control is a moderate fit to the data. Per figure 11, the gaps between Aurora's property crime rate and synthetic Aurora's property crime rates are less than 500 until the treatment year. Similar to Denver's property crime synthetic control, Aurora's property crime synthetic control suffers from opposite trends; while the synthetic control's property crime rates are steadily decreasing every year from 2011, Aurora's property crime rates remain steady until its increase in 2020. However, Aurora's predictors are well-matched by the synthetic control, and Aurora's pretreatment MSPE of 36839.829 is substantially lower than Denver's property crime pretreatment MSPE. Overall, the fit with the data is strong enough to derive meaningful insights.

As we expected, Aurora's property crime trends in 2020 and 2021 roughly follow that of Denver. Aurora similarly had a spike in property crimes that was not matched by the synthetic control. Aurora's property crime rates post-treatment are far above the synthetic control. Some of the gap may be explained by the presence of opposite trends, but the magnitude of the gap makes it plausible that some systematic causal factor is at play.

Table 14: Weights of Top 20 Highly Weighted Jurisdictions

Weights	Unit Names
0.303	Clarksville, Tennessee
0.241	Irving, Texas
0.191	Chesapeake, Virginia
0.085	San Antonio, Texas
0.018	Grand Prairie, Texas
0.011	Round Rock, Texas
0.009	Pasadena, Texas
0.008	Detroit, Michigan
0.008	Houston, Texas
0.007	Odessa, Texas
0.004	Columbia, Missouri
0.003	Frisco, Texas
0.003	Laredo, Texas
0.003	League City, Texas
0.003	Memphis, Tennessee
0.003	Milwaukee, Wisconsin
0.003	Virginia Beach, VA

0.003	Waco, Texas
0.002	Ann Arbor, Michigan
0.002	Arlington, Texas

Results of Placebo Analysis: After removing jurisdictions with pre-treatment MSPEs more than five times greater than that of Aurora, the MSPE ratio of Aurora is statistically significant at both the 10% and 5% levels. Similar to Denver, Aurora's MSPE ratio of 32.703 is surpassed by only Madison, WI and McAllen, TX. The data provides sufficient evidence to indicate that Aurora's 2020 and 2021 property crime rates were significantly greater than those of similar jurisdictions. The significance of the data makes it unlikely that chance alone can explain the increase in property crime rates.

Visualizing the Placebos for Property Crime

Below, we created a histogram to visualize the placebos for property crime and displayed summary statistics.

Table 15: Summary Statistics of Property Crime Placebos

Mean	Standard Deviation	Median
3.918	6.743	1.138

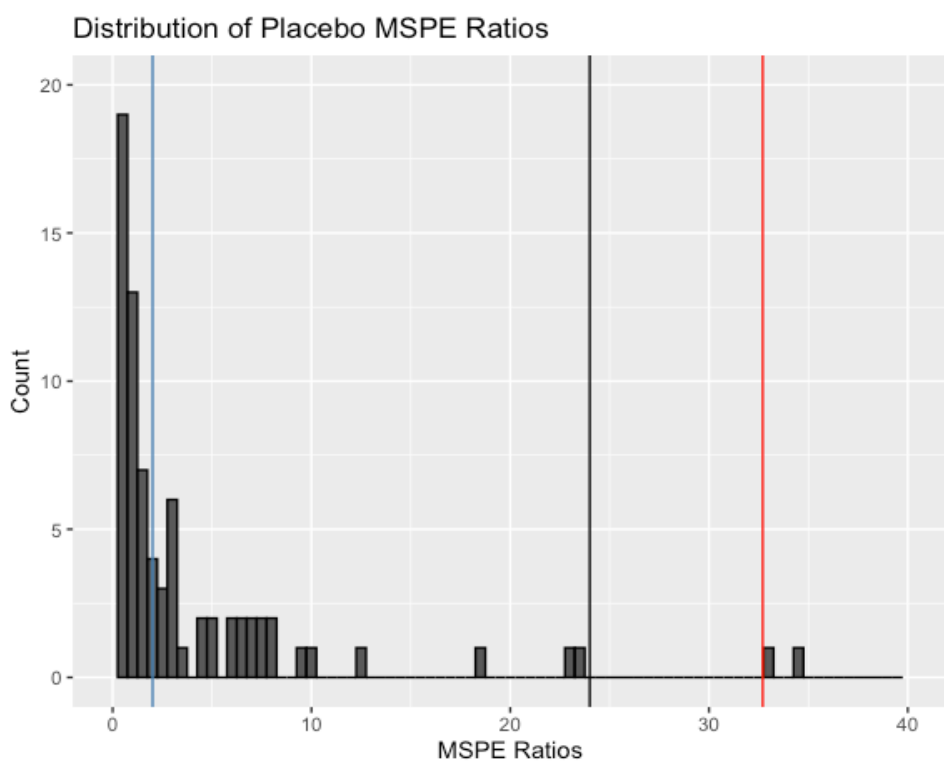


Fig. 12: Placebo MSPE Ratios for Property Crime Rates: The black line represents Denver's MSPE ratio, the red line represents Aurora's MSPE ratio, and the blue line represents Colorado Springs' MSPE ratio. We calculate p-values by dividing the number of "more extreme" MSPE ratios (to the right of the lines) by the total number of MSPE ratios.

Unlike the violent crime MSPE ratios, we did not have any extreme outliers. The standard deviation of the MSPE ratios for the placebos is much lower than the standard deviation of the MSPE ratios for the violent crime placebos even when removing outliers. This implies that the placebo MSPE ratios for property crimes may be more reliable, since there were not many outlier placebo jurisdictions with excessively strong pretreatment fits coupled with significant errors post-treatment.

As we can see on the histogram, both Denver and Aurora's MSPE ratios are extreme compared to the placebo synthetic controls, implying that the property crime increases in both jurisdictions likely did not result purely from chance. On the other hand, Colorado Springs' MSPE ratio is not extreme, existing roughly at the center of the distribution. The implications of this on the hypothesis that Colorado's police accountability law substantially increased property crime rates are mixed at best.

Sensitivity Testing: Displacing by Time

Another method to determine the significance of our results is to change the time of treatment. If changing the inputted treatment time also results in significant results when placebo testing, such a result may indicate that shifts in crime rate from the causal factor at play in 2020 were not significantly larger than shifts in crime rate from past causal factors. In other words, if we can *recreate* the unusually high MSPE ratios of Denver and Aurora in a placebo treatment year, then the shifts created by the real treatment wouldn't be particularly unusual.

We tested the robustness of our model by moving the treatment date to 2017. The post-treatment period was then designated as 2017-2019, and the pretreatment period was designated as 2011-2016. We generated 89 placebos and 3 treatment synthetic controls and calculated MSPE ratios to determine extremity for all 3 jurisdictions. If the model is robust in its result that an unusual 2020 systematic causal factor is at play in Colorado, we would expect generally nonsignificant results in all 3 jurisdictions.

We visualized the placebo distributions below.

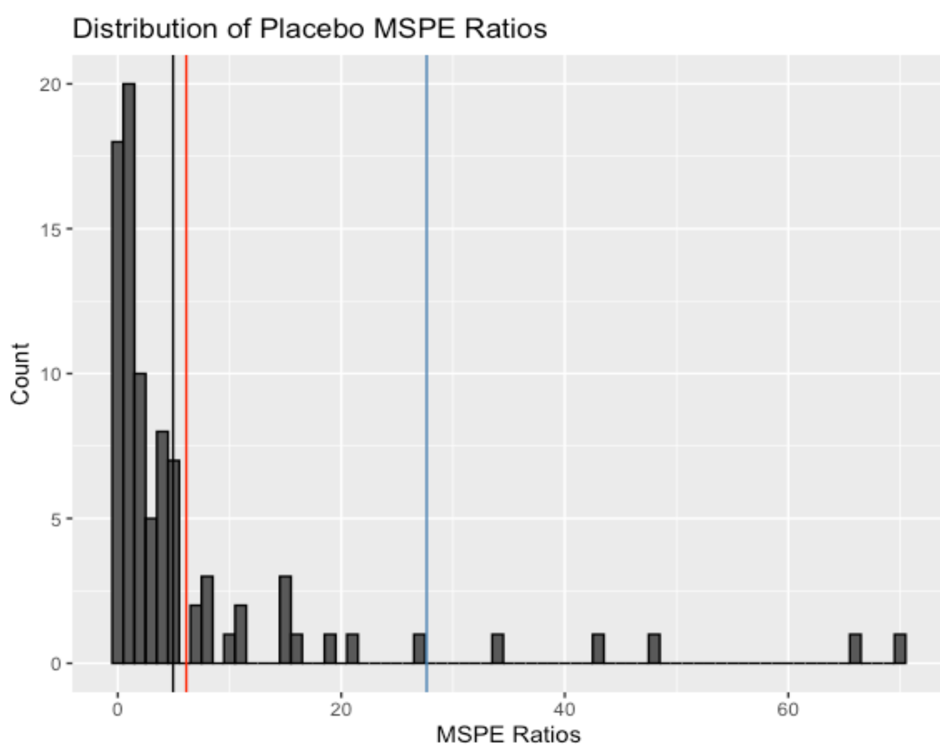


Fig. 13: Placebo MSPE Ratios for Property Crime Rates with Treatment Year 2017: The black line represents Denver's MSPE ratio, the red line represents Aurora's MSPE ratio, and the blue line represents Colorado Springs' MSPE ratio. We calculate p-values by dividing the number of "more extreme" MSPE ratios (to the right of the lines) by the total number of MSPE ratios.

As the histogram demonstrates, the distribution of placebos is affected by a series of outliers above MSPE ratios of 20. This likely results from the fact that we have fewer pretreatment units which skews the pre-treatment MSPE towards lower numbers, creating the possibility for inflated MSPE ratios. This partially explains why Colorado Springs' MSPE ratio appears somewhat extreme; Colorado Springs had an exceptionally strong fit pretreatment in this model before experiencing some deviation in both directions after the placebo treatment time of 2017.

Regardless of its flaws, the histogram demonstrates the robustness of our significant results. None of the three jurisdictions had unusually high MSPE ratios when undergoing a placebo treatment. The exceptionally large post-treatment gaps that we saw in Denver and Aurora were unique to 2020; we could not recreate the effects through placebo treatment years.

Table 16: Comparing Treatment Results with Placebo Treatment (T = treatment year)

	Pre-Treatment MSPE	MSPE Ratio	P-value
T (2020)			
Denver	153872.327	23.9991	0.0253**
Colorado Springs	54693.661	2.0095	0.4521
Aurora	36839.829	32.7034	0.0294**
T – 3 (2017)			
Denver	94999.554	4.9523	0.3289
Colorado Springs	5529.759	27.6425	0.1176
Aurora	41598.403	6.1348	0.2899

* Significant at 10% level

** Significant at 5% level

*** Significant at 1% level

Overall Property Crime Results

We reject our second hypothesis. After constructing 88 different placebos and three synthetic controls for each of the treated jurisdictions, we found that Denver and Aurora both experienced property crime increases significantly greater than those of similar jurisdictions and that such increases likely did not result purely from chance. On the other hand, we found that Colorado Springs' MSPE ratio was not extreme. Therefore, the data provides evidence of a local causal factor in the Denver-Aurora-Lakewood MSA but does *not* provide evidence of a statewide causal factor. We discuss this further in the "Discussion" section.

Summary Table

Table 17: Overall Summary Table

	Pre-Treatment MSPE	MSPE Ratio	P-value
Violent Crime			
Denver	2288.114	14.0741	0.1111
Colorado Springs	1477.022	1.0659	0.6957
Aurora	13362.431	10.0482	0.1190
Property Crime			
Denver	153872.327	23.9991	0.0253**
Colorado Springs	54693.661	2.0095	0.4521
Aurora	36839.829	32.7034	0.0294**

* Significant at 10% level

** Significant at 5% level

*** Significant at 1% level

Discussion

We found no statistically significant evidence in favor of the conclusion that Denver, Colorado Springs, or Aurora experienced unusually high violent crime rates after the passage of the police accountability bill in 2020 compared to control jurisdictions. Although all three jurisdictions did have higher violent crime rates than synthetic controls (to varying degrees), the jurisdictions' MSPE ratios were not extreme when compared with placebos. We do not have enough evidence to say that these jurisdictions' violent crime rates could not have resulted from chance or factors unrelated to the police accountability reform.

On the other hand, we did find statistically significant evidence in favor of the conclusion that Denver and Aurora experienced unusually high property crime rates in 2020 and 2021 compared to control jurisdictions. In particular, Denver and Aurora's property crime rates increased in 2020 and 2021 to be far above the synthetic control, and their calculated MSPE ratios were unusual even in the context of placebos, decreasing the likelihood that chance was the explanation for the property crime increase. We now detail the implications of this result.

Several factors may cast doubt on the property crime findings. First, because both the Denver and Aurora synthetic controls were trending the opposite direction from the observed property crime rates, such synthetic controls are only expected to continue decreasing in 2020 and 2021. The fact that Denver and Aurora experienced large MSPE ratios may simply represent a flaw in the synthetic control itself, not a representation that Denver and Aurora possessed higher property crime rates than expected.

We believe that this concern, although valid, should not invalidate our Denver and Aurora findings. While synthetic Denver and Aurora did trend opposite from the observed cities, they still matched the predictors for both cities extremely well. Standardizing MSPE ratios by dividing by pre-treatment MSPE to account for models that are not well-fit should be able to compensate for some of the error. Additionally, the gaps between the synthetic jurisdictions and the observed jurisdictions were relatively moderate before 2020 at least in the case of Aurora and only expanded dramatically after 2020 and 2021. The synthetic controls, although not fully parallel to observed trends, still imply that some unique causal factor is driving up property crime rates in Aurora and Denver that is not influencing other jurisdictions (or at least not to the same extent).

Second, the predictors we utilized were imperfect. When constructing linear models relating the predictors with the response variables, the predictors for violent crime only had an R-squared value of approximately 0.56, while the predictors for property crime only had an R-squared value of 0.449. In other words, the predictors we chose could only explain roughly 56% of the variation in violent crime rates between jurisdictions and 44.9% of the variation in property crime rates. Because these predictors could not explain significant proportions of the variation in crime rates, synthetic jurisdictions created based on these predictors were imperfect as well. Once again, this concern is valid but should not be enough to discredit the analysis. Especially for phenomenon that is as variable as crime rates, we must accept significant imperfection in choosing the predictors to explain jurisdictional and yearly variations. Although data on certain predictors may improve the analysis (such as data on trust in police), the predictors that we have ensure that the synthetic controls will mimic the real jurisdictions in enough key socioeconomic indicators for the two to possess at least marginally similar crime dynamics.

We are confident in our ability to generate valid, albeit flawed, insights from the methodology that we used. However, we do not believe definitive causal conclusions can be generated from our report.

The data does suggest that some causal factor is uniquely affecting the Denver-Aurora metropolitan statistical area in a way that other control jurisdictions are not experiencing. The data also suggests that such a causal factor likely became prominent in 2020. However, because of the limitations of our synthetic control methodology, we cannot pinpoint the causal factors that explain such an increase. The problem is especially exacerbated given the random variations and unknowns of 2020, ranging from COVID-19 to the George Floyd protests. For instance, it is plausible that Denver and Aurora's policy responses to the COVID-19 pandemic were weaker than surrounding jurisdictions, leading to higher unemployment and more property crimes. It is also plausible that the metropolitan area simply experienced more property destruction during the protests compared to other jurisdictions, leading to more reported property crimes.

Even if we eventually gather the data needed to remove these "unknown" lurking variables, other confounding variables hinder our ability to make an effective causal judgment. It is fully possible that the factors leading to the passage of the police accountability measure also led to increased property crime rates. For instance, citizen distrust of police officers could lead to increased crime rates through decreased cooperation between communities and police. At the same time, citizen distrust could have also generated the political momentum to pass the police accountability reform in the first place. With the presence of all these different plausible causal chains, using these insights to create a definitive claim on what causal factor caused increased property crime in the Denver MSA would be both improper and invalid. For causal inference to be valid, we need more than just statistics; we require all plausible explanatory factors to be controlled for and social scientific evidence that a causal chain is plausible. That is beyond the scope of this report, which only provides the statistics and is unable to control for plausible 2020-2021 explanatory factors.

With these limitations in mind, what are our insights useful for? We believe that our insights can inform the public debate on qualified immunity and police accountability in two ways:

First, although the debate on police accountability usually centers on the effects of such bills on violent crime rates (such as murders and aggravated assaults), our report indicates that there is no significant evidence in favor of the idea that the qualified immunity legislation in Colorado coincided with a greater-than-expected increase in violent crime. In fact, especially in the case of Colorado Springs, the jurisdiction experienced changes in violent crime that were relatively middle of the road compared to placebos. Although the lack of significance does *not* entail that there truly is no relationship between the two variables, we find it unlikely that the Colorado police accountability measure substantially increased violent crime rates in large jurisdictions given that all three jurisdictions did not experience statistically significant increases. In terms of police accountability, our insights suggest that the true evidentiary debate should center on property crime rates.

Second, we can conclude that some unique causal factor increased property crime rates in Denver and Aurora. Our analysis rules out the idea that there is no surge in property crime rate in the MSA; it also rules out chance as the explanation behind the increase. However, the fact that Colorado Springs did not experience a similar level of property crime increase decreases the likelihood that a statewide causal factor, like qualified immunity reform, is the explanation behind such a property crime increase. As we explained in the "Methodology" section, if a statewide causal factor explains the property crime increase, we should see parallel increases across jurisdictions in Colorado, not just in the Denver-Aurora MSA. Colorado Springs did not experience comparable increases to the Denver-Aurora MSA, implying that causal factors unique to the Denver-Aurora MSA caused the increase in property crimes. Nonetheless, at least in those two cities, we believe that our analysis reveals future directions for statistical and social scientific research in determining why those cities experienced such increases.

In sum, we find no statistically significant evidence that qualified immunity reform caused violent crime increases in any of the three jurisdictions we studied. Although we did find evidence of a systematic property increase in Denver and Aurora, the fact that we did not find comparable evidence in Colorado Springs makes it unlikely that a statewide causal factor, such as the police accountability reform, caused the increase. Our report does not *rule out* qualified immunity reform as a causal factor in crime increases or decreases, but we believe our report contributes important evidence as to the plausible effects of police accountability reform on crime rates.

References

Abadie, Alberto. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature* 59, no. 2 (2021): 391-425. doi: 10.1257/jel.20191450.

Amiri, Farnoush. "George Floyd Killing Prompts Some States to Limit or Ban Chokeholds." WKYC Studios, May 23, 2021. https://www.wkyc.com/article/news/nationworld/george-floyd-chokehold-bans/507-3cdf238f-8909-4594-b717-f07eae636e18.

Austin Police Department. *Crime Reports*. Updated July 26, 2021. Accessed July 18, 2021. https://data.austintexas.gov/Public-Safety/Crime-Reports/fdj4-gpfu.

Burgener, Sarah and Champaign Police Department. *2016-2021 Incident Data*. July 13, 2021. NOTE: Data received from FOIA request, not publicly available.

Colorado Congress. Senate. *Concerning Measures to Enhance Law Enforcement Integrity, and, in Connection Therewith, Making an Appropriation (Enhance Law Enforcement Integrity Act)*. SB 20-217. Passed June 19, 2020. http://leg.colorado.gov/sites/default/files/2020a_217_signed.pdf.

Colorado Municipal League. "Requirements of SB 20-217." https://www.cml.org/docs/default-source/uploadedfiles/legislative/requirements-of-sb-20-217--enhance-law-enforcement-integrity-(1).pdf?sfvrsn=ed086c01_2.

Colorado Springs Police Department. *Crime Level Data*. Updated July 2021. Accessed July 2, 2021. https://policedata.coloradosprings.gov/Crime/Crimes-Against-People/ghs7-nqyk.

City and County of Denver and Denver Police Department/Data Analysis Unit. *Crime*. Updated July 28, 2021. Accessed July 2, 2021. https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-crime.

Crime Justice Information Services Division. "Property Crime." 2019 Crime in the United States. Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/property-crime.

Crime Justice Information Services Division. "Table 8: Offenses Known to Law Enforcement by State by City" in 2011 Crime in the United States. 2011. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/tables/table_8_offenses_known_to_law_enforcement_by_state_by_city_2011.xls/view.

Crime Justice Information Services Division. "Table 8: Offenses Known to Law Enforcement by State by City" in 2012 Crime in the United States. 2012. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2012/crime-in-the-u.s.-2012/tables/8tabledatadecpdf/table_8_offenses_known_to_law_enforcement_by_state_by_city_2012.xls/view.

Crime Justice Information Services Division. "Table 8: Offenses Known to Law Enforcement by State by City" in 2013 Crime in the United States. 2013. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2013/crime-in-the-u.s.-2013/tables/table-8/table_8_offenses_known_to_law_enforcement_by_state_by_city_2013.xls/view.

Crime Justice Information Services Division. “Table 8: Offenses Known to Law Enforcement by State by City” in 2014 Crime in the United States. 2014. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2014/crime-in-the-u.s.-2014/tables/table-8/table_8_offenses_known_to_law_enforcement_by_state_by_city_2014.xls/view.

Crime Justice Information Services Division. “Table 8: Offenses Known to Law Enforcement by State by City” in 2015 Crime in the United States. 2015. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2015/crime-in-the-u.s.-2015/tables/table-8/table_8_offenses_known_to_law_enforcement_by_state_by_city_2015.xls/view.

Crime Justice Information Services Division. “Table 8: Offenses Known to Law Enforcement by State by City” in 2016 Crime in the United States. 2016. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016/topic-pages/tables/table-8/table-8.xls/view.

Crime Justice Information Services Division. “Table 8: Offenses Known to Law Enforcement by State by City” in 2017 Crime in the United States. 2017. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017/topic-pages/tables/table-8/table-8.xls/view.

Crime Justice Information Services Division. “Table 8: Offenses Known to Law Enforcement by State by City” in 2018 Crime in the United States. 2018. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/topic-pages/tables/table-8/table-8.xls/view.

Crime Justice Information Services Division. “Table 8: Offenses Known to Law Enforcement by State by City” in 2019 Crime in the United States. 2019. Published by Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/tables/table-8/table-8.xls/view.

Crime Justice Information Services Division. “Violent Crime.” 2019 Crime in the United States. Federal Bureau of Investigation. https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/violent-crime.

Cunningham, Scott. *Causal Inference: The Mixtape*. Yale University Press, 2021. See esp. chap. 10, “Synthetic Control.” https://mixtape.scunning.com/synthetic-control.html.

Equal Justice Initiative. “New Mexico Ends Qualified Immunity for Abusive Police.” Last updated 04/09/2021. https://eji.org/news/new-mexico-ends-qualified-immunity-for-abusive-police/#:~:text=The%20governor%20of%20New%20Mexico,the%20obstacle%20of%20qualified%20immunity.

Federal Bureau of Investigation. “Crime Incident-Based Data by State” in Crime Data Explorer:

Documents and Downloads. https://crime-data-explorer.fr.cloud.gov/pages/downloads.

Federal Bureau of Investigation. "Quarterly Uniform Crime Report" in Crime Data Explorer. https://crime-data-explorer.app.cloud.gov/pages/explorer/crime/quarterly.

Federal Bureau of Investigation. "Crime in the United States Annual Reports" in Crime Data Explorer: Documents and Downloads. https://crime-data-explorer.app.cloud.gov/pages/downloads.

Hill, Anna and Fort Smith Police Department. No title. July 9, 2021. NOTE: Data received from FOIA request, not publicly available.

Houston Police Department. *Monthly Crime Data by Street and Police Beat*. Updated July 2021. Accessed July 26, 2021. https://www.houstontx.gov/police/cs/Monthly_Crime_Data_by_Street_and_Police_Beat.htm.

Maciag, Mike. "Population Density for U.S. Cities." GOVERNING, last updated November 2017. https://www.governing.com/archive/population-density-land-area-cities-map.html.

Seattle Police Department. *SPD Crime Data: 2008-Present*. Updated July 29, 2021. Accessed July 2, 2021. https://data.seattle.gov/Public-Safety/SPD-Crime-Data-2008-Present/tazs-3rd5

Sobel, Nathaniel. "What Is Qualified Immunity, and What Does It Have to Do with Police Reform?" Lawfare Blog. June 6, 2020. https://www.lawfareblog.com/what-qualified-immunity-and-what-does-it-have-do-police-reform

U.S. Census Bureau; American Community Survey. *2011-2019: ACS 5-Year Estimates Data Profiles*. "ACS Demographic and Housing Estimates." https://data.census.gov/cedsci/.

U.S. Census Bureau; American Community Survey. *2011-2019: ACS 5-Year Estimates Data Profiles*. "Selected Economic Characteristics in the United States." https://data.census.gov/cedsci/.

U.S. Census Bureau; American Community Survey. *2011-2019: ACS 5-Year Estimates Data Profiles*. "Selected Housing Characteristics in the United States." https://data.census.gov/cedsci/.

U.S. Census Bureau; American Community Survey. *2011-2019: ACS 5-Year Estimates Data Profiles*. "Selected Social Characteristics in the United States." https://data.census.gov/cedsci/.

Wells, L. E. and Ralph A. Weisheit. "Explaining Crime in Metropolitan and Non-Metropolitan Communities." *International Journal of Rural Criminology* 1, no. 2 (2012): 153-183.

https://doi.org/10.18061/1811/53700.

Appendix 1: Methodology B Results

Methodology B has the advantage of being precise about the treatment dates, allowing us to possibly isolate the qualified immunity bill as a factor, instead of other lurking variables such as the COVID-19 pandemic. However, we do not include the results in our main analysis due to critical methodological limitations that likely invalidate our results. Before we begin the discussion, please note that methodology B was mostly performed in the early stages of the report and has not been revised since. Thus, synthetic control results and inclusion of jurisdictions were much different in methodology B compared to Methodology A.

Violent Crime Synthetic Controls

Synthetic controls were different in several key ways: First, we included all jurisdictions above 50,000 in population with data from all 9 years (2011-2019). Second, we incorporated single female-led household percentage as a predictor and simply used single female-led household percentage with children for 2019. Third, we did not include population density as a predictor. In total, we had around 500 jurisdictions in our synthetic control. Fourth, due to random errors in the optimization functions of the synthetic controls, we varied the pre-treatment time periods to be 2011-2018 and 2011-2019, experimenting with both until one of the functions worked. Such a condition should be largely unimportant, as we do not use the synthetic control to directly match the crime rates, only to determine the series of jurisdictions that, when combined, comprise the majority weight of the synthetic control. Additionally, we optimized over 2014-2019. Below, we display the results for our violent crime synthetic controls.

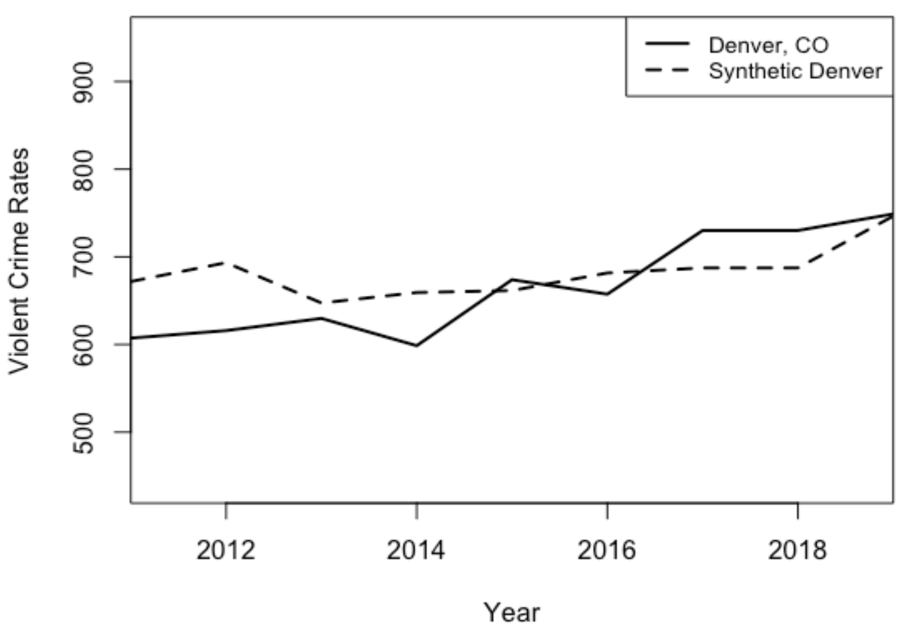


Fig. 14: Path Plot Comparing Denver and Synthetic Denver Violent Crime Rates: The fit is relatively strong throughout all years.

Table 18: Weights of Top 10 Jurisdictions

Weights	Unit Names
0.217	Seattle, Washington
0.206	Houston, Texas
0.199	Fort Smith, Arkansas
0.139	Champaign, Illinois
0.045	Ann Arbor, Michigan
0.033	Des Plaines, Illinois
0.032	Dearborn, Michigan
0.032	Redmond, Washington
0.031	Miami Beach, Florida
0.019	Milpitas, California

The synthetic control fits relatively well for the optimization time period with a pre-treatment MSPE of 1344.333. Using the given weights, we took the five cities with the greatest weights and submitted FOIA requests to obtain access to their incident-level data (if the data was not already public) from 2019-2021. Our request to Ann Arbor was denied, leaving us with a total of four jurisdictions with data. We reran the synthetic control with just those four jurisdictions to determine the jurisdictions' weights for manual calculation of daily violent crime increases. The weights are displayed in the next section.

Property Crime Synthetic Controls

We performed the same method for property crime, except we changed the optimization to 2016-2019 to account for errors when we attempted to run 2014-2019. Because we are not calculating MSPE ratios, it is appropriate to decrease the pre-treatment range to minimize MSPE values over, as an excessively small pre-treatment MSPE does not have disparate impacts on MSPE ratios as they would in a placebo analysis.

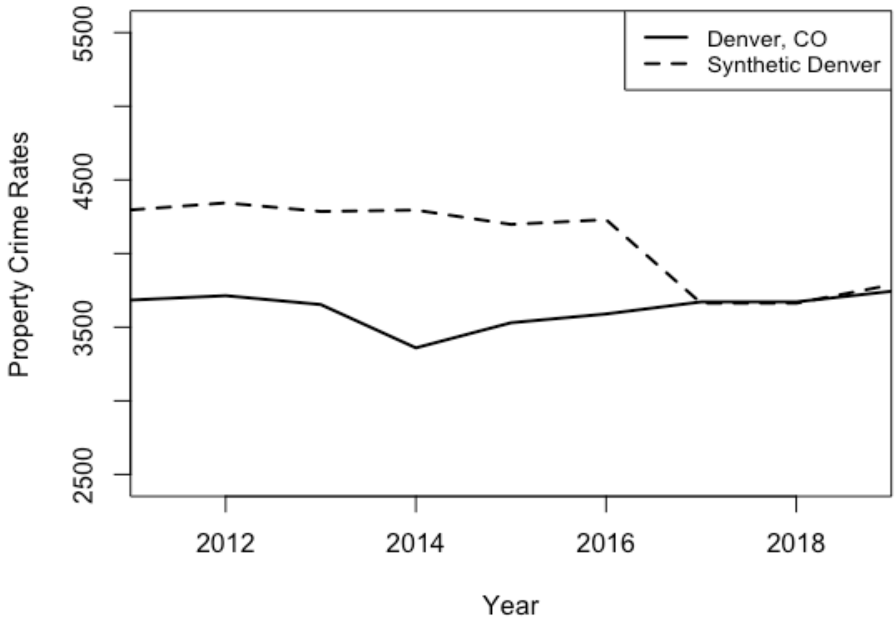


Fig. 15: Path Plot Comparing Denver and Synthetic Denver Property Crime Rates: The fit is extremely weak prior to 2017, where the synthetic control dips to match Denver's rates. This may imply that the control is not very reliable.

Table 19: Weights of Top 10 Jurisdictions

Weights	Unit Names
0.225	Seattle, Washington
0.197	Champaign, Illinois
0.189	Houston, Texas
0.163	Fort Smith, Arkansas
0.042	Austin, Texas
0.028	Cathedral City, California
0.028	Dearborn, Michigan
0.027	Des Plaines, Illinois
0.017	Santa Ana, California
0.015	Milpitas, California

As the path plot demonstrates, synthetic Denver does not follow observed Denver's trends very well, particularly before 2017. Although the pre-treatment MSPE value of 28879.851 appears low, it is important to note that the pre-treatment MSPE is only calculated over the short optimization time period (2016-2019), where the model performs exceptionally well. Nonetheless, we proceeded with the analysis. We once again took the five cities with the highest weights and recalculated the synthetic control. We were able to obtain data from all five jurisdictions.

Statistical Bootstrapping Simulations

After collecting data from each of the four to five 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 proportional daily differences and joined the two datasets together. We used bootstrapping to create a null distribution of 10,000 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:

Table 20: Synthetic Control Weights for Denver Violent Crimes

Name	Weight
Seattle, Washington	0.496466442
Fort Smith, Arkansas	0.487715311
Champaign, Illinois	0.011537542
Houston, Texas	0.004280705

We generated the following two hypotheses:

$H_0: \mu_{Denver} = \mu_{Synthetic}$. 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 equal to the true mean daily proportional difference in number of violent offenses between the two time periods in the synthetic control.

$H_A: \mu_{Denver} > \mu_{Synthetic}$. 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.

$\alpha = 0.05$

Although we had 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. We bootstrapped 10,000 differences in mean assuming no true difference in mean between Denver increases and synthetic control increases and graphically depicted the null distribution below.

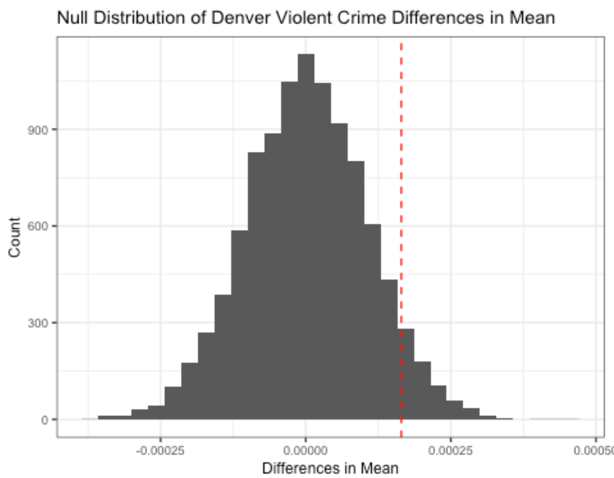


Fig. 16: Null Distribution of Denver Violent Crime Differences in Mean with Control: Each observation in the histogram represents a single simulated difference in mean between Denver and the synthetic control. The red dotted line refers to the observed difference in mean. We took all observations at the observed value or greater and divided by the total number of simulated values to arrive at the p-value.

Because the p-value of 0.0592 is greater than a reasonable alpha level of 0.05, 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 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 for 2019-2021, we decreased the influence of daily crime fluctuations on the results while simultaneously retaining sufficient data points to draw some statistical conclusions.

We created the dummy variables of "time" and "treated" for this end. "Time" takes the value of 1 after June 19, 2020 in both the synthetic control and Denver (with June 20-30 falling under the June 1 value due to monthly numbers), representing the passage of the police accountability legislation. "Treated" takes the value of 1 for Denver and 0 for the synthetic control, representing the jurisdiction designations. The linear model is shown below:

Table 21: Difference in Difference Test for Denver Violent Crimes

Term	Estimate	Std. Error	Statistic	P-value
(Intercept)	390.732	13.913	28.084	0.000

time1	18.503	21.998	0.841	0.404
treated1	-15.177	19.676	-0.771	0.444
time1:treated1	37.942	31.110	1.220	0.228

The interaction variable for variables "time" and "treated" reflects the difference-in-difference estimate. Because the p-value of 0.228 far exceeds a reasonable alpha level of 0.05, we fail to reject the null hypothesis. The data does not provide sufficient evidence on the monthly level that the passage of qualified immunity reform on June 19, 2020 in Denver corresponded 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:

Table 22: Synthetic Control Weights for Denver Property Crimes

NAME	weight
Austin, Texas	0.6935988
Champaign, Illinois	0.2188792
Seattle, Washington	0.0874752
Fort Smith, Arkansas	0.0000465
Houston, Texas	0.0000002

Since Houston had a negligible weight, we decided to exclude Houston from the analysis and run the bootstrapping with data from the other four jurisdictions.

$H_0: \mu_{Denver} = \mu_{Synthetic}$. 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 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 greater than the true mean daily proportional difference in number of violent offenses between the two time periods in the synthetic control model.

$\alpha = 0.05$

Once again, we bootstrapped 10000 differences in mean, assuming that the null hypothesis is true. We graphically depicted the null distribution below:

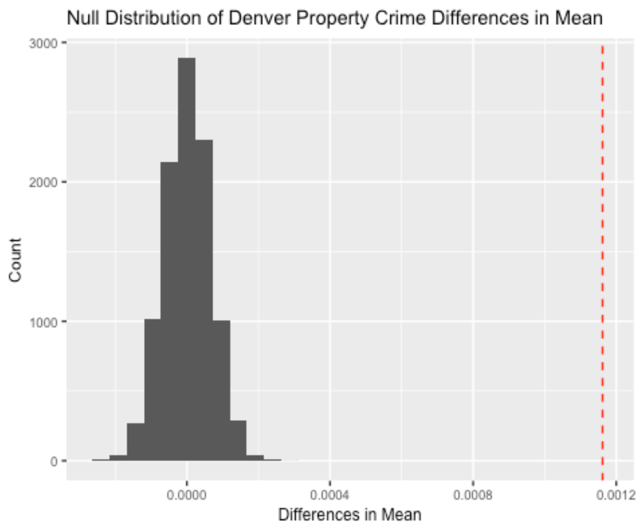


Fig. 17: Null Distribution of Denver Property Crime Differences in Mean with Control: Each observation in the histogram represents a single simulated difference in mean between Denver and the synthetic control. The red dotted line refers to the observed difference in mean. We took all observations at the observed value or greater and divided by the total number of simulated values to arrive at the p-value. This time, the observed value far surpasses any of the values of the null distribution.

Because the p-value of 0 is far less than a reasonable alpha level of 0.05, 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. The graph comparing the synthetic control trends with the true Denver trends is not parallel, especially from 2016-2017 (see Figure 14).

Summary of Results

This methodology concludes similarly to the previous methodology we used. The data does not provide evidence at the 1% or 5% significance levels to indicate that Denver's increase in violent crime after the passage of police accountability legislation 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. Because of the critical limitations in our data and Methodology, these results do not meet the standard of statistical rigor needed to present this as definitive evidence that property crime rates truly did increase in Denver beyond what was expected. For instance, substantial problems existed in the way that we simulated to obtain results. By using daily differences between two different time periods, the standard deviation of such differences were exaggerated, as crimes can randomly increase or decrease day by day without reference to broader legislation. If there happened to be 20 violent crime incidents on June 19, 2020 and 0 violent crime incidents on June 19, 2019, the methodology would flag that day as a highly significant violent crime increase, even though the two days are not interconnected in any way. Additionally, our method of standardization gave smaller jurisdictions disproportionately more weight, as tiny variations in violent crime incidents were far more significant. The graph below displays this phenomenon visually.

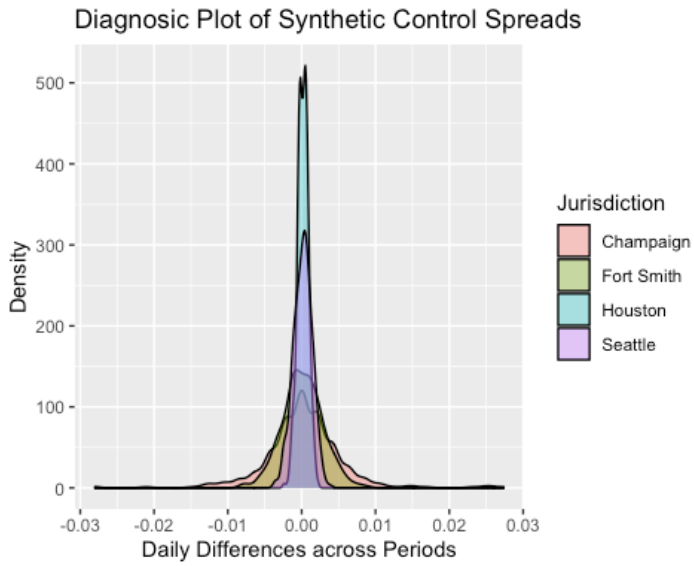


Fig. 18: Density Plot of Synthetic Control Standardized Differences in Violent Crime: Each of the numbers on the x-axis represents the difference between the number of violent crimes on a 2020-21 day and a 2019-20 day divided by the total number of violent crimes in the 2019-20 period. Noticeably, the spreads of each jurisdiction are correlated with their respective populations.

Nonetheless, we include the methodology here to demonstrate possible conclusions of an analysis that accurately referenced the treatment date and to provide additional corroboration of the main findings of our report.

Appendix 2: Investigating Outliers

As noted in the "Results" section, the placebo synthetic controls for violent crime were significantly skewed by a series of high-MSPE ratio outliers, including two extreme outliers with MSPE ratios greater than 80. We investigate these outliers now.

Manchester, NH

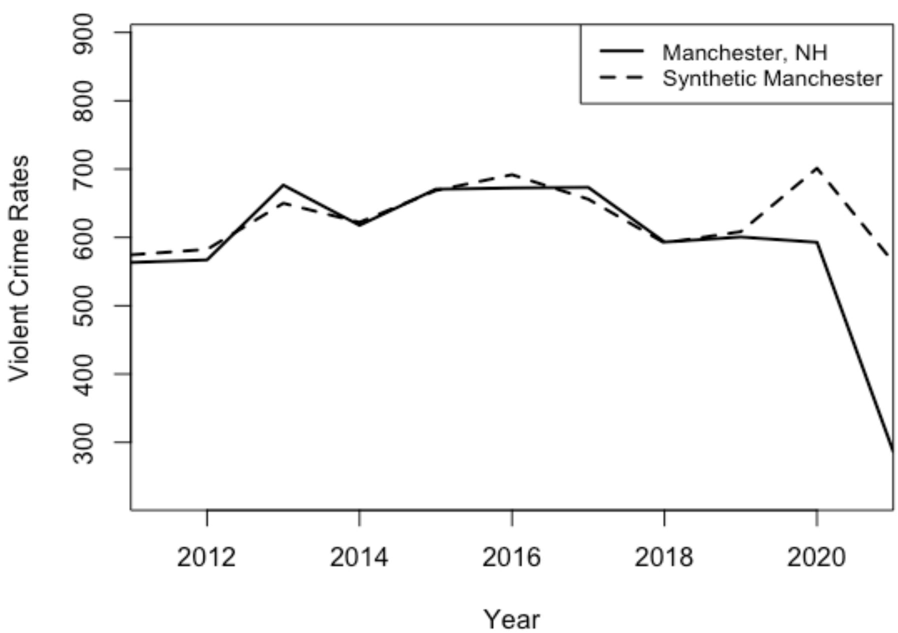


Fig. 19: Path Plot Comparing Manchester and Control Violent Crime: While the synthetic control follows observed Manchester exceptionally well until 2019, Manchester experiences a decrease in violent crime rate that is not followed by the control in 2020 and 2021.

Manchester, NH had an MSPE ratio of over 200, more than 40 times greater than the mean MSPE ratio without Manchester or Evansville. The path plot provides some insight into the mathematical reasons behind this occurrence. The synthetic control tracks the violent crime trends in Manchester exceptionally well until 2020; in 2020, the synthetic control experiences an increase in violent crime that is not matched by Manchester itself. Additionally, Manchester's violent crime rate in the first 3 quarters of 2021 is far lower than expected by the synthetic control. Thus, Manchester had an extremely low pre-treatment MSPE coupled with a large post-treatment MSPE. Although this is an outlier, further investigation does not reveal any clear data errors or differing circumstances that would warrant removing the Manchester data from the dataset. Likely, this resulted from the weakness of our predictors coupled with our inability to track unknowns in 2020 and 2021; plausibly, Manchester had a stronger response to the COVID-19 pandemic or less police distrust that allowed it to avoid the violent crime increases that the rest of the country faced.

Evansville, Indiana

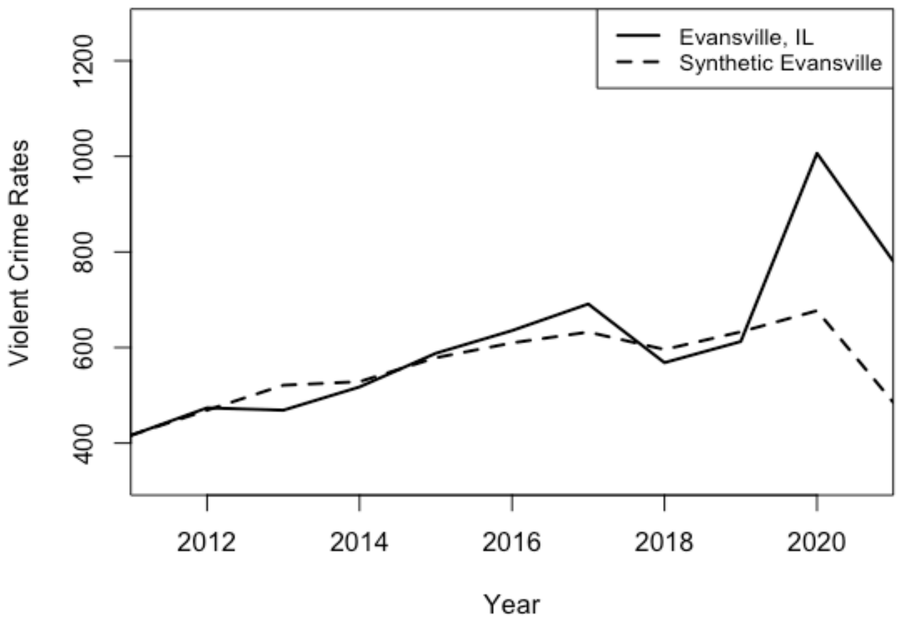


Fig. 20: Path Plot Comparing Evansville and Control Violent Crime: While the synthetic control follows observed Evansville exceptionally well until 2019, Evansville experiences an increase in violent crime rate that is not followed by the control in 2020 and 2021.

Similarly, in Evansville, the synthetic control tracks the violent crime trends well until the treatment period begins, where Evansville experiences a sharp increase in violent crimes that is not followed by the synthetic control. Based on these graphs, we can make 2 plausible conclusions:

First, the synthetic control method appears to have trouble tracking abrupt shifts in violent crime rates, which may be attributable to the randomness by which violent crime rates increase or decrease. In Aurora's violent crime synthetic control, Aurora's abrupt increase in violent crime rates in 2016 was not well-tracked either, implying that although the synthetic control is effective at following trends over time, outlier years cannot be accurately followed with the predictors that we have. Similarly, in Evansville and Manchester, sharp increases and decreases in violent crime rates even without treatment could not be successfully tracked by the synthetic control.

Second, it is plausible that, even without treatment, there could be large increases in violent crime rates that simply happen to fall on the post-treatment years. This highlights the difficulty of making a causal claim; because there are many lurking variables, and violent crime rates are often very random phenomena, we cannot attribute large increases in crime rates purely to treatments. We can, however, use significance testing to diminish the likelihood that the years are explainable purely by chance, as we do in the analysis.

Whether these conclusions apply to property crime analysis is less certain, as the property crime placebos did not have many significant outliers.

Exhibit D

National Police Accountability Project
A Project of the National Lawyers Guild

IMPACT OF THE NEW MEXICO CIVIL RIGHTS ACT ONE YEAR LATER

In July 2022, NPAP polled fifteen members who practice civil rights in New Mexico to tell us how the New Mexico Civil Right Act (NMCRA) has impacted the rights of their clients since going into effect in July 2021. Overall, NPAP attorneys have not filed more cases than usual but they believe NMCRA will help ensure that their clients who suffered constitutional violations will not have their cases dismissed or stalled because of qualified immunity.

- **NM Civil Rights Attorneys Are Not Filing Significantly More Cases.**
    - Most attorneys that had sued under the law responded that they added NMCRA claims to cases they would have otherwise filed as a standard Section 1983 action.
    - Only two members reported filing a case exclusively under NMCRA and not Section 1983.
    - Five members reported having a case in development that they think would be vulnerable to dismissal under qualified immunity if it were filed as a standard Section 1983 case but will survive under NMCRA.
- **NMCRA Will Help Civil Rights Plaintiffs Survive Dispositive Motions.**
    - Most members anticipate that NMCRA will help them survive dispositive motions on qualified immunity.
    - Members are also optimistic that NMCRA will help their cases against institutional defendants since it creates a cause of action against them, as well. In *Hand v. Cty. of Taos, NM*, the District court found the plaintiff had stated a claim against the county board under NMCRA but not Section 1983 because he had not identified an official policy or custom in his complaint. 2022 U.S. Dist. LEXIS 115462, at *5-6 (D.N.M. June 29, 2022).
- **NMCRA Is Helping Avoid Delays Associated with Qualified Immunity**
    - One member thinks that the NMCRA will also lower the number of motions to dismiss, interlocutory appeals, and discovery stays caused by qualified immunity. He is basing this on the fact that a defendant he regularly sues did not file a motion to dismiss in a case where he added NMCRA claims.

2022 St. Bernard Avenue, Suite 310 | New Orleans, Louisiana 70116
504-220-0401 tel | www.npapjustice.org