

# The Effect of Three Strikes Laws on Violent Crime Rates

Lonnie Hofmann\*

October 20, 2020

## Abstract

Three strikes laws mandate sharply increased sentences for criminals that commit a specific number of felonies. I analyze the effect of these laws on violent crime rates using municipal-level data from the FBI. Specifically, I compare violent crime rates of border municipalities in states with differing treatment statuses using a difference-in-differences specification with a sample matched on pre-treatment outcomes. I find that three strikes laws do not reduce violent crime rates. I rule out reductions in violent crime rates greater than 1.3%, and I reject the hypothesis that three strikes laws reduce violent crime rates at the 5% significance level. Additional analyses and robustness checks support my main findings.

*Keywords:* difference-in-differences, matching

## 1 Introduction

Three strikes laws, a type of habitual offender law passed by many states in the 1990s, aim to reduce violent crime rates by locking up repeat, violent offenders for longer periods of time. In addition, they use the threat of enhanced prison sentences to deter criminals from committing violent crimes. However, the effectiveness of three strikes laws in reducing violent crime rates has been unclear.

My paper contributes to the existing literature in a few ways. First, many studies have analyzed the effects of a three strikes law passed in California in 1994. However, fewer studies have examined how three strikes laws impacted crime outcomes in other states, and fewer

---

\*Department of Economics, University of Missouri. E-mail: [lrh6df@mail.missouri.edu](mailto:lrh6df@mail.missouri.edu)

still have studied crime outcomes in other states using county or municipal-level data. As such, there exist real contributions to be made to the three strikes literature by examining the effect of three strikes laws on crime outcomes at the local level. My paper makes such a contribution by analyzing the impact of three strikes laws at the municipal level in all states that passed such a law.

Furthermore, this paper improves on previous identification strategies used to analyze three strikes laws by restricting the analysis to within “border regions”. In other words, I restrict my sample to include only “border municipalities”; i.e., municipalities located in a county that borders another state. I take inspiration from Card and Krueger’s (1993) seminal minimum-wage study that evaluated the effect of a New Jersey minimum wage law on employment using a sample of fast food restaurants along the Pennsylvania/New Jersey border. Other papers have used similar techniques. For example, Dube, Lester, and Reich (2010) study the effects of minimum wages on employment using a sample of contiguous county pairs across state borders.

Using the aforementioned border region restriction and a sample matched on pre-treatment outcomes, I analyze the effect of three strikes laws on violent crime rates using a municipal-level difference-in-differences specification. I find no statistical evidence that three strikes laws decrease violent crime. I rule out reductions in violent crime rates greater than 1.3%, and I reject the hypothesis that three strikes laws reduce violent crime rates at the 5% significance level. Additional analyses and robustness checks support my main findings.

Section 2 provides a literature review and background information on three strikes laws. Section 3 describes the data. Section 4 discusses the empirical strategy. Section 5 outlines the main results. Section 6 suggests directions for future research.

## 2 Three Strikes Laws

### 2.1 Literature Review

Many papers have examined the effect of California's 1994 three strikes law on crime outcomes as California's law was one of the most far-reaching in the nation (Clark, Henry, and Austin, 1997). In this literature, the results are mixed. Various papers have found that the California law deterred crime (Datta, 2017; Marcet, 2011; Shepherd, 2002), reduced criminal activity (Iyengar, 2008), reduced felony arrest rates (Helland and Tabarrok, 2007), or had an incapacitation effect on instrumental crimes (Ramirez and Crano, 2003).<sup>1</sup> Conversely, other papers find no evidence that the California law had a deterrent or incapacitating effect on crime (Worrall, 2004), or reduced crime rates below their expected level (Stolzenberg and D'alessio, 1997).

In addition, a few studies have analyzed three strikes laws using state-level data. Chen (2008) finds that three strikes laws reduced crime. On the other hand, Marvell and Moody (2001) find that three strikes laws increased homicides.

While many of the California studies conduct analysis at the county or municipal level, fewer papers evaluate three strikes laws in multiple states using such data. One such paper is Kovandzic, Sloan III, and Vieraitis (2004) who analyze the effect of three strikes laws in 188 cities with populations of at least 100,000. They find that three strikes laws increased homicide rates but had no effect on other crime rates.

### 2.2 Law Details

During the 1990s, twenty-four U.S. states passed some type of three strikes law (Clark, Henry, and Austin, 1997). These laws mandate increased sentences for criminals that commit two to four qualifying felonies (the exact number depends on the state). The status of three

---

<sup>1</sup>Instrumental crime, as defined in Ramirez and Crano (2003), includes illegal acts such as robbery, burglary, and motor vehicle theft.

strikes laws for all U.S. states is summarized in Figure 1. States that are orange, blue, or yellow passed a three strikes law in the 1990s. States in grey did not pass a three strikes law during that time.

The Federal Government also passed a national three strikes law in 1994 (Clark, Henry, and Austin, 1997). However, violations of federal law can only be tried in federal court. In addition, the jurisdiction of federal courts is not as broad as that of state courts (The Federal Judicial Center, 2006). Thus, most violent crimes of interest (murder, robbery, rape, aggravated assault, etc.) fall under the purview of state courts, and are subject to state-level three strikes laws (The Federal Judicial Center, 2006).

The crimes that constitute a “strike” vary among states, but there are some constants. For example, most violent crimes (murder, rape, robbery, aggravated assault, etc.) constitute a strike (Clark, Henry, and Austin, 1997). Some states count other crimes as strikes as well. For example, selling drugs to minors is a strike in California (Clark, Henry, and Austin, 1997).

Furthermore, the consequences of repeat offending differ from state to state. In twelve states, repeat offenders receive a mandatory life sentence without the possibility of parole (Clark, Henry, and Austin, 1997). In other states, repeat offenders may receive a life sentence, but with the possibility of parole (Clark, Henry, and Austin, 1997). Regardless, in all twenty-four states that passed a three strikes law, enhanced sentencing (in some form) occurs for repeat violent offenders (Clark, Henry, and Austin, 1997).

In addition, the number of strikes required to trigger enhanced sentencing varies in a few states. For most states, a criminal needs to receive three strikes to be eligible for enhanced sentencing (Clark, Henry, and Austin, 1997). However, Arkansas, Georgia, Montana, North Dakota, South Carolina, and Tennessee require only two strikes to trigger enhanced sentencing (Clark, Henry, and Austin, 1997). On the other hand, Maryland requires four strikes (Clark, Henry, and Austin, 1997).<sup>2</sup>

---

<sup>2</sup>There are subtleties regarding the number of strikes required to trigger enhanced sentencing in many states. For example, Arkansas requires only two strikes to trigger enhanced sentencing for the most serious

Prior to the passage of three strikes legislation in the 1990s, all 50 states (and the District of Columbia) had existing habitual offender laws that allowed for enhanced sentencing of repeat criminal offenders (Woodard, 1991).<sup>3</sup> Therefore, the results in this paper should be interpreted as the effects of enhancing penalties conditional on existing habitual offender laws. For example, prior to passing three strikes legislation in 1994, Indiana allowed for adding 8–30 years on top of normal sentences for habitual felony offenders (Woodard, 1991). However, after passing three strikes legislation in 1994, Indiana enhanced existing penalties so that a criminal who commits a third violent felony receives a mandatory life sentence without parole (Clark, Henry, and Austin, 1997).

Three strikes laws in 18 of 24 states focus solely on repeat, violent offenders (Clark, Henry, and Austin, 1997).<sup>4</sup> Since such offenders would be most affected by passage of three strikes laws, I restrict the outcome variables in my analyses to be measures of violent crime.

### 3 Data

My paper uses violent crime data from Uniform Crime Reporting (UCR). UCR is compiled by the FBI and gives official data on reported criminal offenses in the United States. The data I use originates from municipalities with at least 10,000 residents and covers the time period from 1988–2000.

As defined by UCR, violent crime consists of four offenses: murder and non-negligent manslaughter, rape, robbery, and aggravated assault. UCR has precise definitions for what constitutes each offense, but reporting inconsistencies can happen. For example, numerous

---

of offenses (such as murder), but it requires three strikes for less serious offenses (such as robbery) (Clark, Henry, and Austin, 1997). Clark, Henry, and Austin (1997) outline further legislative specifics for all 24 states that passed three strikes laws.

<sup>3</sup>The literature disagrees on whether Kansas had a habitual offender law that allowed for enhanced sentencing for repeat criminals prior to the passage of its three strikes law in 1994. Woodard (1991) claims Kansas had such a law; Clark, Henry, and Austin (1997) claim it did not. I choose to follow Woodard (1991). Either way, the existence or non-existence of Kansas’s pre-1994 habitual offender law does not substantially affect the analyses in this paper.

<sup>4</sup>California, Florida, Indiana, Louisiana, South Carolina, and Washington are the six states that include some non-violent crimes as “strikes” (Clark, Henry, and Austin, 1997).

municipalities reported rape offenses incorrectly to UCR during the 1985–2014 time period. Consequently, UCR removed rape and violent crime counts from its database in the years where there was a reporting issue for the municipalities in question.<sup>5</sup> Thus, in order to retain as many municipalities in my analysis as possible, I construct my own version of violent crime. Since the reporting issues mostly affect rape data, my adjusted violent crime variable consists of the three other offenses, besides rape, that constitute the UCR definition of violent crime (i.e., murder and non-negligent manslaughter, robbery, and aggravated assault). So, henceforth when I refer to violent crime, I refer to my adjusted version of violent crime that omits rape data.<sup>6</sup>

My analyses also incorporate regressors that are deemed predictive of crime by the crime literature. Municipal-level regressors are not available for all municipalities in my sample, so I use county-level regressors instead. These county-level regressors include male share of the population, black share of the population, Hispanic share of the population, share of the population between 18 and 44 years old, share of the population without a high school diploma, share of the population in poverty, and unemployed share of the population. The regressor data comes from samples of U.S. census data compiled by IPUMS USA (Flood, Goeken, Grover, Meyer, Pacas, Ruggles, and Sobek, 2020).

The regressor dataset from IPUMS does not include data for every year of my panel, so I linearly interpolate to fill in the data for the years that are missing. Additionally, for any county present in the UCR data but not present in the census data, I set its regressor values to be the mean of regressor values from counties present in both data sets. Thus, the regressor values for counties appearing only in UCR do not change over time. To account for this, I include a “missing regressor” fixed effect in my regressions which equals one if a particular county lacks regressor data. This allows me to retain my full sample while ensuring my estimates are only affected by regressor values that change over time.

---

<sup>5</sup>This issue was especially pronounced in Illinois where every single municipality had its rape and violent crime counts removed for the years 1985–2005.

<sup>6</sup>As a robustness check, I run my analyses using violent crime rates that include rape data. This version of the analysis does not change my results substantially. These results are compiled in Appendix B.

Summary statistics for my data are listed in Table 1. I list statistics for my full sample in the left-most column, and statistics for both my treated and control sub-samples in the next two columns. The total sample and each sub-sample are fairly similar across all metrics, although my control sub-sample has a slightly higher black share of the population (13.1%) compared to my treated sub-sample (9.6%). Note the high standard deviations for the violent crime rate across all samples (502, 485, and 523 for Total, Treated, and Control, respectively, compared to means of 428, 444, and 406). This indicates a high degree of dispersion in the outcome data.

## 4 Methodology

### 4.1 Main Analysis

My main specification is the following difference-in-differences regression:

$$Y_{mcr t} = \alpha + \beta T_{mt} + \mathbf{X}_{\mathbf{crt}} + \delta_m + \theta_{rt} + \epsilon_{mcr t} \quad (4.1)$$

where  $m$  indexes municipalities,  $r$  indexes border regions,  $t$  indexes years, and  $c$  indexes counties.  $Y_{mcr t}$  is the violent crime rate per 100,000 people, and  $T_{mt}$  is the treatment indicator variable.  $T_{mt}$  equals one in all years including and following the year that the state containing municipality  $m$  passed a three strikes law.  $T_{mt}$  equals zero in all years prior to the state containing municipality  $m$  passing a three strikes law.  $\mathbf{X}_{\mathbf{crt}}$  is vector of county-level regressors that includes male share of the population, black share of the population, Hispanic share of the population, share of the population that is between 18 and 44 years old, share of the population without a high school diploma, share of the population in poverty, and unemployed share of the population.  $\delta_m$  are municipality fixed effects,  $\theta_{rt}$  are border-region by year fixed effects,  $\alpha$  is the intercept, and  $\epsilon_{mcr t}$  is the error term.

Some of the municipalities in the data have undergone changes to their crime reporting structures over time, which can affect measured outcomes. To address this, I specify the

municipality fixed effects as municipality-by-reporting-regime fixed effects. For example, if Mobile, Alabama changed its reporting structure in 1998, pre-1998 Mobile and post-1998 Mobile would be considered separate municipalities in my analysis.

The border region by year fixed effects in (4.1) and my sample selection process are inspired by the analyses of minimum wage laws in both Dube, Lester, and Reich (2010) and Card and Krueger (1993). Card and Krueger (1993) examine a sample of fast-food restaurants on either side of the New Jersey–Pennsylvania border in order to analyze the effect of a New Jersey minimum-wage law on employment. By restricting their sample to only include border region restaurants, Card and Krueger (1993) obtain more accurate findings since any estimated effect is due to differences in law status (based on where the restaurants are located) and not due to differing restaurant characteristics. I hope to accomplish something similar by restricting my sample to only include municipalities located in counties that border another state.

A complication is that some municipalities are located in counties that border multiple states. However, I cannot assign two or more border regions to a given municipality. Thus, these municipalities are assigned the border region that includes the nearest border state with a differing treatment status. For example, let’s consider a municipality in Washington that is located in a county that borders both Idaho and Oregon but is geographically closer to Idaho. In this case, I assign its border region to be Idaho–Washington (IDWA), as opposed to Oregon–Washington (ORWA), since Idaho and Oregon both have different treatment statuses than Washington. On the other hand, let’s consider an Arkansas municipality that is located in a county that borders both Missouri and Tennessee but the municipality is geographically closest to Tennessee. Here, I assign its border region to be Arkansas–Missouri (ARMO), as opposed to Arkansas–Tennessee (ARTN), because Missouri and Arkansas have differing treatment statuses while Arkansas and Tennessee’s treatment statuses are the same.

Once I restrict my sample as stated above, my border-region fixed effects then restrict the analysis further so that I am only comparing municipalities *within* a border region (Dube,



Lester, and Reich, 2010). Thus, my estimated effects of three strikes laws will come from variation within municipalities over time, and variation within a border region and year due to differences in treatment status.

I cluster the confidence intervals from (4.1) at the level of treatment (i.e., the state level). In addition, I use the wild cluster bootstrap to construct the confidence intervals for my parameters of interest because typical cluster-robust standard errors can be inaccurate in analyses with a small number of clusters (Cameron, Gelbach, and Miller, 2008).<sup>7</sup> I choose to use the wild cluster bootstrap to construct my confidence intervals out of an abundance of caution even though my analysis exceeds the typical cutoff thought necessary to use conventional cluster-robust standard errors (i.e., 42 clusters as stated in Angrist and Pischke (2008)).

Goodman-Bacon (2018) points out the potential for bias in difference-in-differences specifications due to treatment timing heterogeneity. However, treatment timing occurs near the middle of my study period for all treated states (i.e., during the years 1993–1995, with the full study period ranging from 1988–2000), thus reducing concerns of such bias. In addition, Goodman-Bacon’s (2018) concerns do not apply to event studies. Thus, the fact that my event study results in Section 5 are similar to my results using the difference-in-differences specification provides additional support for minimal bias resulting from treatment timing heterogeneity.

## 4.2 Pre-Trend Analysis

The key underlying assumption for a difference-in-differences model is the parallel trends assumption. To test this assumption, I utilize the following event study specification:

$$Y_{mrt} = \alpha + \beta_5 T_{mt}^5 + \beta_4 T_{mt}^4 + \beta_3 T_{mt}^3 + \beta_2 T_{mt}^2 + \beta_0 T_{mt}^0 + \beta_1 T_{mt}^1 + \beta_2 T_{mt}^2 + \beta_3 T_{mt}^3 + \beta_4 T_{mt}^4 + \beta_5 T_{mt}^5 + \mathbf{X}_{\text{crt}} + \delta_m + \theta_{rt} + \epsilon_{mrt} \quad (4.2)$$

---

<sup>7</sup>Specifically, I use the recommended options in the “boottest” package in Stata to construct my wild cluster bootstrap confidence intervals (Roodman, Nielsen, MacKinnon, and Webb, 2019).

where  $T_{mt}^0$  is an indicator variable set equal to one for the year  $t$  in which the state containing municipality  $m$  passed a three strikes law,  $T_{mt}^5$  is set equal to one five or more years after passage of a three strikes law in the state containing municipality  $m$ ,  $T_{mt}^4$  is set equal to one four years after passage of the law, etc. Similarly,  $T_{mt}^5$  is set equal to one five or more years prior to the passage of a three strikes law in the state containing municipality  $m$ ,  $T_{mt}^4$  is set equal to one four years prior to passage of the law, etc.  $T_{mt}^1$  is omitted from the model in order to serve as a baseline for comparison.  $\mathbf{X}_{\text{crt}}$ ,  $\delta_m$ ,  $\theta_{rt}$ ,  $\alpha$ , and  $\epsilon_{mrt}$  are as stated in Section 4.1. Evidence consistent with the parallel trends assumption holding would consist of the parameter estimates for  $T_{mt}^5$ ,  $T_{mt}^4$ ,  $T_{mt}^3$ , and  $T_{mt}^2$  being close to zero and statistically insignificant.

Unfortunately, my pre-treatment parameter estimates are different from zero and statistically significant when I run (4.2) on my full sample. In other words, there is evidence of poor parallel pre-trends. As a result, I use propensity-score matching on pre-treatment outcomes in order to improve my pre-trends.<sup>8</sup> Specifically, I use 1:N “nearest-neighbor” propensity score matching (with replacement), enforcing common support, with five control units matched to each treatment unit.<sup>9</sup> After the matching process is completed, I run (4.1) and (4.2) on the matched sample.

There is disagreement in the matching literature regarding whether to match on pre-treatment outcomes in order to improve pre-trends for difference-in-differences analyses. Ryan, Kontopantelis, Linden, and Burgess Jr (2019) and Chabé-Ferret (2017) find that matching on pre-treatment outcomes works well in simulations. Other papers have found that matching on pre-treatment outcomes can lead to bias when the pre-treatment outcomes are correlated with treatment status (Daw and Hatfield, 2018). However, this particular issue is not a concern for me as the correlation between pre-treatment outcomes and treatment

---

<sup>8</sup>Matching is done via the “MatchIt” software in R (Ho, Imai, King, and Stuart, 2007; Stuart, King, Imai, and Ho, 2011).

<sup>9</sup>Mueser, Troske, and Gorislavsky (2007) find that matching only one control to each treatment (with replacement) can lead to increased standard errors due to sampling error, whereas using five or ten nearest neighbors does not increase standard errors to the same extent. Consequently, I choose to use five nearest neighbors in an attempt to avoid this sampling error problem.

status is very low in my analysis.<sup>10</sup>

## 5 Results

### 5.1 Border Municipality Analysis

The pre-trend analysis for (4.2) run on my matched sample is shown in Figure 2. Figure 2 includes point estimates and 95% confidence intervals for the estimated treatment effect in each year within an 11-year window around the treatment year (specifically, five years before and after the treatment year). Estimates for the year immediately preceding treatment are omitted in order to serve as a baseline. Y-axis units are a rate per 100,000 people.

Evidence consistent with the parallel trends assumption holding would include the pre-treatment coefficient estimates in (4.2) being close to zero and statistically insignificant. My analysis suffers from being underpowered, as can be seen by the somewhat wide confidence intervals. However, the pre-treatment coefficient estimates in Figure 2 are indeed close to zero and statistically insignificant providing evidence consistent with the parallel trends assumption holding.

Roth (2018) argues that there is potential for bias from pre-testing and many pre-tests are under-powered. While my analysis does suffer from being somewhat under-powered, I reduce the potential for bias in my pre-trend testing since my study period is fairly short (1988-2000). This means there is less time for other events to occur that could bias my event study estimates pre-treatment. In addition, I manually ensure that my treatment and control samples are similar in both trends and levels prior to treatment by matching on pre-treatment outcomes. This also makes the pre-trends assumption more plausible.

My results from running (4.1) on my border municipality sample are compiled in Table 2.<sup>11</sup> Table 2 includes point estimates, 95% wild cluster bootstrap confidence intervals

---

<sup>10</sup>In my sample, the Pearson correlation coefficient equals 0.054, indicating very low correlation between pre-treatment outcomes and treatment status.

<sup>11</sup>Various robustness checks that slightly adjust my main specification are included in Appendix B. My

clustered at the state level, and both two-tailed and lower-tailed p-values.

Columns (1) and (2) contain results for (4.1) run on the non-matched sample. Columns (1) and (2) include border-region weights that give equal weight to each border region that is included in (4.1). Column (2) contains county-level regressors while column (1) does not. While the point estimates for columns (1) and (2) are not statistically significant, this version of my analysis does not satisfy the parallel trends assumption so one should interpret the results with caution.

Columns (3) and (4) in Table 2 show results where municipalities are matched based on pre-treatment outcomes and then (4.1) is run on the matched sample. The matching is done via 1:N “nearest-neighbor” propensity-score matching (with replacement), enforcing common support, with five controls matched to each treatment. The only variables used for the matching are violent crime rates one, two, three, four, and five years prior to treatment.

The point estimates in columns (3) and (4) are statistically significant at the ten percent level for specifications with and without county-level controls. My preferred specification in column (4), with municipality fixed effects, border-region by year fixed effects, county-level controls, and matching weights, indicates that the passage of a three strikes law led to the violent crime rate increasing by 58.9 violent crimes per 100,000 people. This translates to a 14.2% increase in the violent crime rate. However, my analysis is somewhat under-powered in that I cannot rule out the possibility of a negative effect. However, I can rule out a negative effect larger than 1.3%. Consequently, I reject the hypothesis that three strikes laws reduce violent crime rates at the 5% significance level, shown via the p-value for my lower-tailed test in Table 2. Additionally, the fact that that my point estimates do not change significantly when adding in controls (i.e., moving from column (3) to (4)) gives support to my choice of specification as an identification strategy.

Note that the border municipality results in Table 2 include “spillover” effects. That is, if we believe criminals to be rational actors, we might expect passage of a three strikes

---

results are robust to these specification changes.

law in one state to cause criminals to move their criminal activity across the border into a state without a three strikes law. If this movement of criminal activity from treated to non-treated states exists, it would be captured in my border municipality analysis. Furthermore, the direction of bias from any “spillover” effects is likely to be negative. Yet, despite the potential presence of this negative bias, my border municipality results are still positive. In order to disentangle such “spillover” effects from the law’s real effect, I pursue an interior county analysis in Section 5.2.

## 5.2 Interior Municipality Analysis

In order to disentangle the possible effect of criminals moving criminal activity across borders from treated to non-treated states (i.e., “spillover” effects), I analyze the effect of three strikes laws on violent crime rates using a sample of interior municipalities. I again use (4.1) as my specification, but now my sample consists of interior municipalities from states with different treatment statuses. An additional small change is that I omit county-level controls for the interior county analysis since my results in Section 5.1 do not change substantially when they are included.

In order to maintain the same specification across my border municipality and interior municipality analyses, I assign a “border-region” to each interior municipality. This presents a problem as interior municipalities, by definition, are located in counties that do not border other states. To rectify the problem, I randomly assign border regions to interior municipalities based on the frequency that border municipalities appear in a given border region. For example, California borders three states: Oregon, Nevada, and Arizona. This means each California border county has one of three border regions: California-Oregon (CAOR), California-Nevada (CANV), and California-Arizona (CAAZ). For the sake of our example, let’s say the border region distribution for California border municipalities is as follows: 10% of California border municipalities are located in the CAOR border region, 40% in the CANV border region, and 50% in the CAAZ border region. This means, for my interior

county analysis, 10% of interior California municipalities would be randomly assigned the CAOR border region, 40% would be randomly assigned the CANV border region, and 50% would be randomly assigned the CAAZ border region.

The pre-trend analysis for (4.2) run on my interior municipality matched sample, using the same matching procedure as in Section 5.1, is shown in Figure 3. Figure 3 includes point estimates and 95% confidence intervals for the estimated treatment effect in each year within an 11-year window around the treatment year (specifically, five years before and after the treatment year). Estimates for the year immediately preceding treatment are omitted in order to serve as a baseline. Y-axis units are a rate per 100,000 people.

Evidence consistent with the parallel trends assumption holding would include the pre-treatment coefficient estimates in (4.2) being close to zero and statistically insignificant. Similar to the the border municipality analysis, my interior municipality analysis suffers from being under-powered, as can be seen by the somewhat wide confidence intervals. However, the pre-treatment coefficient estimates in Figure 3 are indeed close to zero and statistically insignificant providing evidence consistent with the parallel trends assumption holding.

My results from running (4.1) on my interior municipality sample are compiled in Table 3. Table 3 includes point estimates, 95% wild cluster bootstrap confidence intervals clustered at the state level, and both two-tailed and lower-tailed p-values.

Column (1) contains results for (4.1) run on the non-matched sample. As with my border municipality analysis, Column (1) includes border-region weights that give equal weight to each border region that is included in (4.1). The point estimate for column (1) is statistically significant at the 5% level. However, as this un-matched version of my analysis does not satisfy the parallel trends assumption, one should interpret the results with caution.

Column (2) in Table 3 displays results where I match municipalities based on pre-treatment outcomes and then run (4.1) on the matched sample. As in Section 5.1, the matching is done via “nearest-neighbor” propensity-score matching (with replacement), enforcing common support, with five controls matched to each treatment. The only variables

used for the matching are violent crime rates one, two, three, four, and five years prior to treatment. In addition, the specification in column (2) includes municipality fixed effects, border-region by year fixed effects, and matching weights.

The point estimate in column (2) is statistically significant at the ten percent level and indicates that the passage of a three strikes law led to the violent crime rate increasing by 33.3 violent crimes per 100,000 people. This translates to a 7.4% increase in the violent crime rate. As in my border municipality analysis, my analysis is somewhat under-powered in that I cannot rule out the possibility of a negative effect. However, I can rule out a negative effect larger than 1.7%. Consequently, I reject the hypothesis that three strikes laws reduce violent crime rates at the 5% significance level, shown via the p-value for my lower-tailed test in Table 3.

If the passage of three strikes laws in treatment states caused crime to spill over into non-treatment states, I would expect my estimates in Section 5.1 to overestimate the negative effect of the laws on crime. Thus, the removal of such spillovers, via the interior municipality analysis, would result in less negative estimates. In fact, the results from Section 5.2 are statistically indistinguishable from Section 5.1, providing no evidence of spillover effects.

### 5.3 Falsification Tests

To test whether my results in Section 5.1 are spurious, I evaluate the effect of three strikes laws on outcomes that should be unaffected by passage of such laws. The outcomes used in these “placebo” (or falsification) tests include all the county-level regressors used in my main specification: male population share, black population share, Hispanic population share, share of the population aged 18-44, share of the population in poverty, unemployed share of the population, and share of the population without a high school degree. If a county-level regressor is used as the outcome variable, it is omitted as a control variable. Otherwise, the falsification test specification is identical to the main specification outlined in Section 5.1.

My results from the falsification tests are displayed in Table 4. Table 4 includes point

estimates, 95% wild cluster bootstrap confidence intervals clustered at the state level, and two-tailed p-values.

Each column contains results for a different outcome variable. Six of the seven variables have small point estimates and are statistically insignificant at all common significance levels. The lone exception is black share of the population which is statistically significant at the 5% level. However, the point estimate is still somewhat small (my results indicate that passage of a three strikes law results in a 1.4% increase in the black share of the population), and I can rule out an effect larger than 2.5%. In addition, some statistically significant results are likely to occur by chance when undertaking multiple hypothesis testing. In summary, I believe the falsification tests provide support for my findings in Section 5.1, with the lone statistically significant result for black share of the population likely to have occurred by chance.

## 5.4 Heterogeneity in Three Strikes Sentencing

A potential problem with analyzing how three strikes laws affect crime outcomes is that states applied three strikes sentencing with varying degrees of enthusiasm. For example, according to the Prison Policy Initiative (1998), only six states (California, Florida, Georgia, Nevada, South Carolina, and Washington) sentenced more than 100 criminals using provisions in their respective three strikes legislation in the 2–5 years after passage of said laws. California, specifically, sentenced 26,074 people under its three strikes law as of December 31, 1996 (Clark, Henry, and Austin, 1997), far and away the most of any state that passed a three strikes law.

In Becker’s (1968) seminal work on economics and crime, he models the way in which individual decisions to commit crimes are affected by both the probability of getting caught and the strictness of the punishment if convicted. The latter effect from Becker (1968) is most relevant in the context of heterogeneous three strikes sentencing. In other words, the anticipated deterrent effects (disincentivizing individuals from committing crimes due to



harsh sentences) and incapacitation effects (locking up criminals for long periods of time so they can no longer commit crimes) of three strikes laws may not be fully realized due to infrequent use of enhanced sentencing in many states.

That being said, unless policymakers specifically write stipulations requiring homogeneous and widespread implementation of a law, heterogeneous implementation will be unavoidable. As a result, researchers cannot control how a law is implemented; all they can do is analyze the real-world effects that the data provides. Thus, analyses of a policy will capture implementation heterogeneity in their estimates of the policy's effects. In other words, this paper's analysis of the effects of three strikes laws on violent crime rates provides effect estimates when three strikes implementation is heterogeneous.

In addition, I attempt to evaluate the effect of three strikes laws on crime outcomes in states that applied the law frequently. Specifically, I restrict my sample to the six states that sentenced more than 100 criminals using provisions in their respective three strikes legislation in the 2–5 years after passage of said laws (Prison Policy Initiative, 1998). Unfortunately, my analysis is significantly under-powered due to small sample sizes, so the results are inconclusive. Nevertheless, the results are included in Appendix B for the sake of completeness.

## 5.5 Discussion of Disparate Results in the Three Strikes Literature

In Section 5.1, I find no statistical evidence that three strikes laws reduced violent crime rates. However, there are mixed findings in the three strikes literature. Most of the papers that find three strikes laws reduced crime analyzed California's three strikes law, whereas most that find three strikes laws had no effect on crime studied three strikes laws in multiple states.<sup>12</sup> In addition, most papers that analyzed California's three strikes law found that crime reduction came via a deterrent effect, i.e., harsher penalties deterred criminals (or potential criminals) from committing crimes.

---

<sup>12</sup>Chen (2008) is the exception in that she undertook a multi-state analysis and found three strikes laws reduced crime.

Interestingly, enhanced penalties deterring crime has little support in the deterrent effect literature. For example, many criminals are unaware of the sentences they face when committing a crime (Austin, Eisen, Cullen, Frank, Chettiar, and Brooks, 2017; Wright, 2010). Thus, they do not consider the potential sentence when considering whether to commit a crime (Austin et al., 2017; Wright, 2010). To illustrate, during interviews with numerous male inmates at state prisons in the late 1990's, Anderson (2002) found that a staggering 89% of the most violent criminals either perceived no risk of getting caught or were unaware of the potential penalties for the crimes they committed. In addition, criminals seem to discount future years in prison. So, they do not view enhanced penalties as harshly as authors of three strikes laws might anticipate. Paternoster (2010) references interviews with arrested individuals who perceive a five-year prison sentence to be only twice as severe as a one-year sentence, and who perceive a 20-year sentence to be only one-half times as severe as a 10-year sentence. Furthermore, there is evidence that the probability of being caught has much more of a deterrent effect on crime than the severity of the sentence (Durlauf and Nagin, 2011; Paternoster, 2010). In fact, Antunes and Hunt (1973) point out that severity of punishment deters crime only when the probability of being caught is high. Finally, Wright (2010) points out that as many as half of all state prisoners were under the influence of drugs and/or alcohol at the time they committed the offense for which they were imprisoned. This casts doubt on the assumption in many economic models of crime that criminals are rational actors who weigh the costs and benefits before committing a crime.

In brief, I find that three strikes laws did not reduce crime, and the deterrent effect literature finds minimal evidence that harsh sentences deter crime. As such, I hypothesize that three strikes laws have little to no impact on violent crime rates via a deterrent effect. However, the possibility remains that three strikes laws could reduce crime via incapacitation effects. For example, my panel is probably too short to adequately investigate incapacitation effects. Clark, Henry, and Austin (1997) point out that three strikes laws apply to serious

offenses and such cases take longer to reach disposition.<sup>13</sup> So, incapacitation effects could exist but they have yet to appear in the data. Furthermore, most three strikes papers focus on deterrent effects, so the question of incapacitation effects in the context of three strikes laws remains largely unstudied.<sup>14</sup> On the other hand, as mentioned in Section 5.4, three strikes sentencing was applied infrequently in most states. Consequently, this would limit the scope of any incapacitation effects.

## 6 Conclusion and future work

Using a sample matched on pre-treatment outcomes, I analyze the effect of three strikes laws on violent crime rates using a municipal-level difference-in-differences specification. I find no statistical evidence that three strikes laws reduce violent crime. I rule out reductions in violent crime rates greater than 1.3%, and I reject the hypothesis that three strikes laws reduce violent crime rates at the 5% significance level. Additional analyses and robustness checks support my main findings.

As discussed in Section 5.5, many papers have studied whether three strikes laws deterred crime. However, fewer have studied incapacitation effects. Given that the deterrent effects literature lacks support for enhanced sentencing deterring crime, it would be interesting to see whether three strikes laws reduce crime via incapacitation effects. Additional studies that build on the incapacitation effect work of Ramirez and Crano (2003), Worrall (2004), and Chen (2008) would be welcome.

Furthermore, numerous papers have examined California's three strikes law. Yet, few papers have analyzed three strikes laws in other states that apply enhanced sentencing frequently. Studying three strikes laws in these states would clarify whether frequent application of three strikes sentencing would reduce violent crime rates.

Nonetheless, my paper encourages policymakers to act cautiously when deciding whether

---

<sup>13</sup>Disposition means the final outcome of a prosecution.

<sup>14</sup>Ramirez and Crano (2003), Worrall (2004), and Chen (2008) are the exceptions.

to pass a three strikes law. If policymakers hope to reduce violent crime rates, my paper suggests three strikes laws may not accomplish their goal.

## References

- Anderson, David A. 2002. "The deterrence hypothesis and picking pockets at the pickpocket's hanging." *American Law and Economics Review* 4 (2):295–313.
- Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Antunes, George and A Lee Hunt. 1973. "The impact of certainty and severity of punishment on levels of crime in American states: An extended analysis." *The Journal of Criminal Law and Criminology (1973-)* 64 (4):486–493.
- Austin, James, Lauren-Brooke Eisen, James Cullen, Jonathan Frank, Inimai Chettiar, and Cornell William Brooks. 2017. "How many Americans are unnecessarily incarcerated?" *Federal Sentencing Reporter* 29 (2-3):140–174.
- Becker, Gary S. 1968. "Crime and punishment: An economic approach." In *The Economic Dimensions of Crime*. Springer, 13–68.
- Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics* 90 (3):414–427.
- Card, David and Alan B Krueger. 1993. "Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania." Tech. rep., National Bureau of Economic Research.
- Chabé-Ferret, Sylvain. 2017. "Should we combine difference in differences with conditioning on pre-treatment outcomes?" Tech. rep., TSE Working Paper.
- Chen, Elsa Y. 2008. "Impacts of "three strikes and you're out" on crime trends in California and throughout the United States." *Journal of Contemporary Criminal Justice* 24 (4):345–370.
- Clark, John, D Alan Henry, and James Austin. 1997. *"Three strikes and you're out": A review of state legislation*. US Department of Justice, Office of Justice Programs, National Institute of Justice.
- Datta, Anusua. 2017. "California's three strikes law revisited: Assessing the long-term effects of the law." *Atlantic Economic Journal* 45 (2):225–249.
- Daw, Jamie R and Laura A Hatfield. 2018. "Matching and regression to the mean in difference-in-differences analysis." *Health Services Research* 53 (6):4138–4156.
- Dube, Arindrajit, T William Lester, and Michael Reich. 2010. "Minimum wage effects across state borders: Estimates using contiguous counties." *The Review of Economics and Statistics* 92 (4):945–964.
- Durlauf, Steven N and Daniel S Nagin. 2011. "Imprisonment and crime: Can both be reduced?" *Criminology & Public Policy* 10 (1):13–54.
- Flood, Sarah, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, Steven Ruggles, and Matthew Sobek. 2020. "IPUMS USA: Version 10.0 [dataset]." *Minneapolis, MN: IPUMS, 2020* URL <https://doi.org/10.18128/D010.V10.0>.

- Goodman-Bacon, Andrew. 2018. “Difference-in-differences with variation in treatment timing.” Tech. rep., National Bureau of Economic Research.
- Helland, Eric and Alexander Tabarrok. 2007. “Does three strikes deter? A nonparametric estimation.” *Journal of Human Resources* 42 (2):309–330.
- Ho, Daniel E, Kosuke Imai, Gary King, and Elizabeth A Stuart. 2007. “Matching as non-parametric preprocessing for reducing model dependence in parametric causal inference.” *Political Analysis* 15 (3):199–236.
- Iyengar, Radha. 2008. “I’d rather be hanged for a sheep than a lamb: The unintended consequences of ‘three-strikes’ laws.” Tech. rep., National Bureau of Economic Research.
- Kovandzic, Tomislav V, John J Sloan III, and Lynne M Vieraitis. 2004. ““Striking out” as crime reduction policy: The impact of “three strikes” laws on crime rates in US cities.” *Justice Quarterly* 21 (2):207–239.
- Marcet, Daniel. 2011. “Three strikes and you’re out: A triple differences approach to estimating the deterrent effect of California’s three strikes law.” Unpublished Honors Thesis.
- Marvell, Thomas B and Carlisle E Moody. 2001. “The lethal effects of three-strikes laws.” *The Journal of Legal Studies* 30 (1):89–106.
- Mueser, Peter R, Kenneth R Troske, and Alexey Gorislavsky. 2007. “Using state administrative data to measure program performance.” *The Review of Economics and Statistics* 89 (4):761–783.
- Paternoster, Raymond. 2010. “How much do we really know about criminal deterrence?” *The Journal of Criminal Law and Criminology* :765–824.
- Prison Policy Initiative. 1998. ““Three strikes” laws: Five years later.” URL <https://static.prisonpolicy.org/scans/sp/3strikes.pdf>.
- Ramirez, Juan R and William D Crano. 2003. “Deterrence and incapacitation: An interrupted time-series analysis of California’s three-strikes law.” *Journal of Applied Social Psychology* 33 (1):110–144.
- Roodman, David, Morten Ørregaard Nielsen, James G MacKinnon, and Matthew D Webb. 2019. “Fast and wild: Bootstrap inference in Stata using boottest.” *The Stata Journal* 19 (1):4–60.
- Roth, Jonathan. 2018. “Pre-test with caution: Event-study estimates after testing for parallel trends.” Tech. rep., Working Paper.
- Ryan, Andrew M, Evangelos Kontopantelis, Ariel Linden, and James F Burgess Jr. 2019. “Now trending: Coping with non-parallel trends in difference-in-differences analysis.” *Statistical Methods in Medical Research* 28 (12):3697–3711.
- Shepherd, Joanna M. 2002. “Fear of the first strike: The full deterrent effect of California’s two- and three-strikes legislation.” *The Journal of Legal Studies* 31 (1):159–201.
- Stolzenberg, Lisa and Stewart J D’alessio. 1997. ““Three strikes and you’re out”: The impact of California’s new mandatory sentencing law on serious crime rates.” *Crime & Delinquency* 43 (4):457–469.
- Stuart, Elizabeth A, Gary King, Kosuke Imai, and Daniel Ho. 2011. “MatchIt: Nonparametric preprocessing for parametric causal inference.” *Journal of Statistical Software* .
- The Federal Judicial Center. 2006. “Federal courts & what they do.” URL <https://ar.usembassy.gov/wp-content/uploads/sites/26/2016/03/FCtsWh06.pdf>.
- Woodard, Paul L. 1991. *Statutes requiring the use of criminal history record information*. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

- Worrall, John L. 2004. “The effect of three-strikes legislation on serious crime in California.” *Journal of Criminal Justice* 32 (4):283–296.
- Wright, Valerie. 2010. *Deterrence in criminal justice: Evaluating certainty vs. severity of punishment*. Sentencing Project.

## A Matching Summary Statistics

Additional summary statistics for my sample after matching are included in Table 5.

## B Robustness Checks

Table 6 includes results from running (4.1) on a matched sample where I match on pre-treatment outcomes by setting the treatment year for control states to be the average treatment year across the entire sample. Also, note that the matching procedure is the same as in Table 2 except for the additional use of a caliper set equal to 0.05.<sup>15</sup> As can be seen in Table 6, my results are robust to this specification change.

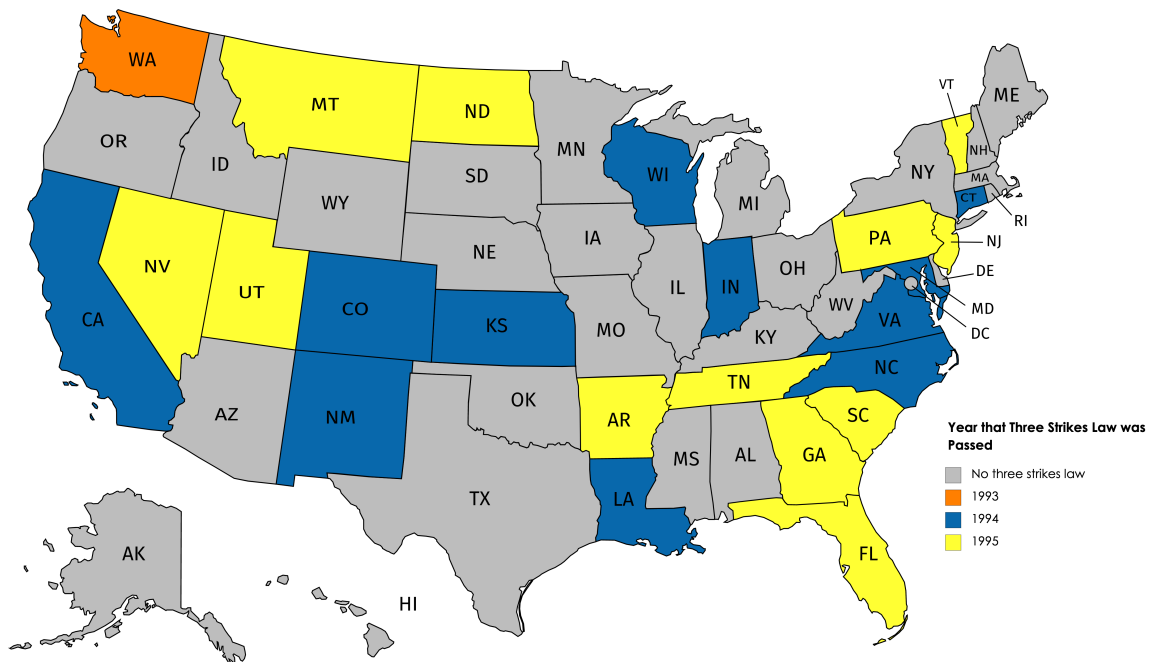
Table 7 includes results from running (4.1) on a matched sample where only municipalities from border regions where one state is a treatment state and the other state is a control state are included. In other words, I no longer include municipalities from border regions with two treatment states that have differing treatment timings. As can be seen in Table 7, my results are robust to this specification change.

Table 8 includes results from running (4.1) on a matched sample where I use an unadjusted measure of the violent crime rate as the outcome variable. Specifically, this measure of the violent crime rate includes rapes in addition to murder and non-negligent manslaughter, robbery, and aggravated assault. As can be seen in Table 8, my results are robust to this specification change.

---

<sup>15</sup>Per Stuart et al. (2011, pg. 19), the caliper is defined as “the number of standard deviations of the distance measure within which to draw control units”.

Table 9 includes results from running (4.1) on a matched sample where the sample is restricted to the six states (California, Florida, Georgia, Nevada, South Carolina, and Washington) who sentenced more than 100 criminals using provisions in their respective three strikes legislation in the 2-5 years after passage of said laws (Prison Policy Initiative, 1998). Unfortunately, my analysis is significantly under-powered due to small sample sizes, so the results are inconclusive.



Created with mapchart.net

Figure 1: States that passed a three strikes law and the year of passage



Table 1: Summary Statistics

	Total Mean (SD)	Treated Mean (SD)	Control Mean (SD)
Violent Crime	428 (502)	444 (485)	406 (523)
Male	48.8 (0.8)	48.8 (0.8)	48.7 (0.8)
Black	11.3 (8.1)	9.6 (7.5)	13.1 (8.3)
Hispanic	11.1 (9.7)	11.9 (11.1)	10.2 (8.0)
18-44 years old	40.9 (2.6)	40.8 (2.8)	40.9 (2.4)
No HS degree	38.5 (4.4)	38.8 (5.0)	38.2 (3.6)
In poverty	10.3 (3.7)	9.9 (3.8)	10.7 (3.6)
Unemployed	6.2 (1.7)	6.0 (1.7)	6.4 (1.6)
N	12,103	6,305	5,798

Violent Crime is a rate (per 100,000 people). All other variables are in terms of share of the population (out of 100). N is the number of municipality-year observations. Violent crime data comes from Uniform Crime Reporting (UCR); all other data comes from Flood et al. (2020).

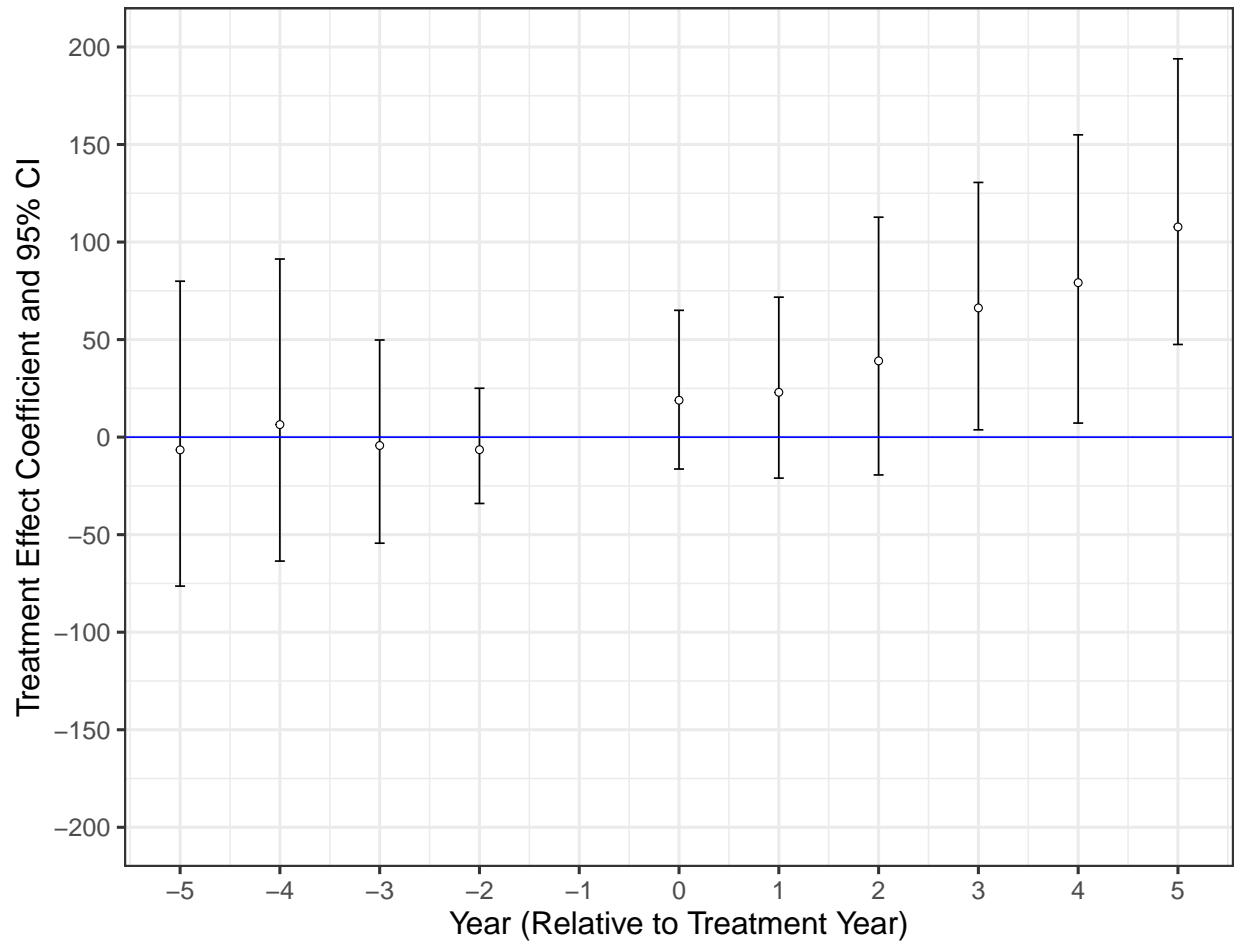


Figure 2: Border Municipality Pre-Trend Analysis

Table 2: Effects of Three Strikes Laws on Violent Crime Rates, Border Municipality Analysis

	(1)	(2)	(3)	(4)
policy	20.3 [-28.7 , 71.5]	26.6 [-27.0 , 83.7]	64.9 [-8.2 , 125.7]	58.9 [-5.3 , 122.4]
Municipality FE	X	X	X	X
Border-Region x Year FE	X	X	X	X
County Controls		X		X
Border-Region Weights	X	X		
Matching Weights			X	X
$N$	12,103	12,103	9,373	9,373
Mean Dependent Variable	428	428	414	414
No. of clusters	46	46	43	43
P-value ( $H_0 : \beta = 0$ )	0.439	0.384	0.090*	0.065*
P-value ( $H_0 : \beta < 0$ )	0.224	0.188	0.049**	0.029**

Notes: \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a three strikes law and is in terms of a rate per 100,000 people. Wild cluster bootstrap confidence intervals clustered at the state level are in square brackets and are calculated using 1,000 replications. County controls include male population share, black population share, Hispanic population share, share of population aged 18-44, share of population in poverty, share of population that is unemployed, and share of population without a high school degree.  $N$  is the number of municipality-year observations. Matching is done via 1:N “nearest-neighbor” propensity score matching (with replacement), enforcing common support, with five controls matched to each treatment.

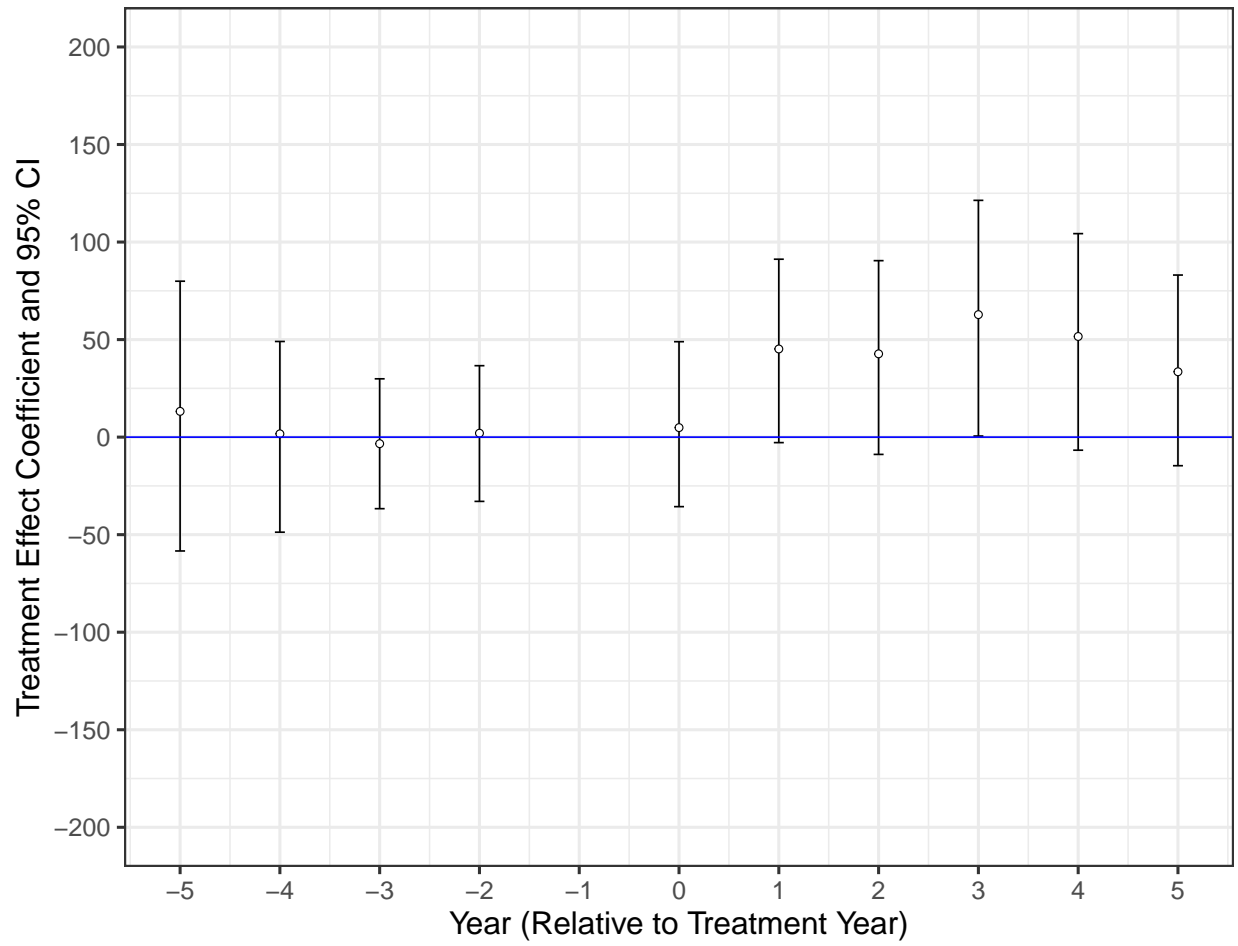


Figure 3: Interior Municipality Pre-Trend Analysis

Table 3: Effects of Three Strikes Laws on Violent Crime Rates, Interior Municipality Analysis

	(1)	(2)
policy	59.7 [1.9 , 124.6]	33.3 [-7.8 , 74.2]
Municipality FE	X	X
Border-Region x Timing FE	X	X
County Controls		
Border-Region Weights	X	
Matching Weights		X
$N$	25,519	21,307
Mean Dependent Variable	476	448
No. of clusters	42	42
P-value ( $H_0 : \beta = 0$ )	0.043**	0.095*
P-value ( $H_0 : \beta < 0$ )	0.024**	0.044**

Notes: \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a three strikes law and is in terms of a rate per 100,000 people. Wild cluster bootstrap confidence intervals clustered at the state level are in square brackets and are calculated using 1,000 replications. County controls include male population share, black population share, Hispanic population share, share of population aged 18-44, share of population in poverty, share of population that is unemployed, and share of population without a high school degree.  $N$  is the number of municipality-year observations. Matching is done via 1:N “nearest-neighbor” propensity score matching (with replacement), enforcing common support, with five controls matched to each treatment.

Table 4: Effects of Three Strikes Laws on Other Outcomes

	Male	Black	Hispanic	Young	No HS	Poverty	Unemployed
policy	-0.07 [-0.17 , 0.05]	0.16 [0.02 , 0.28]	0.13 [-0.47 , 0.71]	0.08 [-0.67 , 0.60]	-0.40 [-0.78 , 0.10]	-0.08 [-0.25 , 0.12]	-0.05 [-0.18 , 0.06]
Municipality FE	X	X	X	X	X	X	X
Border-Region x Year FE	X	X	X	X	X	X	X
County Controls	X	X	X	X	X	X	X
Border-Region Weights							
Matching Weights	X	X	X	X	X	X	X
$N$	9,373	9,373	9,373	9,373	9,373	9,373	9,373
Mean Dependent Variable	48.7	11.1	11.0	41.0	38.4	10.1	6.1
No. of clusters	43	43	43	43	43	43	43
P-value ( $H_0 : \beta = 0$ )	0.214	0.024**	0.633	0.739	0.103	0.335	0.314

Notes: \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates the treatment effect for the variable indicated in the column. Policy is in terms of share of the population (out of 100). Wild cluster bootstrap confidence intervals clustered at the state level are in square brackets and are calculated using 1,000 replications. County controls include male population share, black population share, Hispanic population share, share of population aged 18-44, share of population in poverty, share of population that is unemployed, and share of population without a high school degree. A variable is omitted as a control when it is used as an outcome variable.  $N$  is the number of municipality-year observations. Matching is done via 1:N “nearest-neighbor” propensity score matching (with replacement), enforcing common support, with five controls matched to each treatment. The variables used for matching are violent crime rates one, two, three, four, and five years prior to treatment.

Table 5: Summary Statistics (Matching)

	Total Mean (SD)	Treated Mean (SD)	Control Mean (SD)
Violent Crime Rate	414 (495)	426 (465)	395 (538)
Male	48.7 (0.8)	48.8 (0.8)	48.6 (0.8)
Black	11.1 (8.3)	9.4 (7.4)	13.2 (8.8)
Hispanic	11.0 (9.8)	11.9 (11.0)	9.7 (7.8)
18-44 years old	41.0 (2.7)	40.9 (2.8)	41.2 (2.4)
No HS degree	38.4 (4.5)	38.6 (5.1)	38.1 (3.6)
In poverty	10.1 (3.8)	9.6 (3.8)	10.8 (3.8)
Unemployed	6.1 (1.7)	5.9 (1.7)	6.3 (1.6)
N	9,373	5,291	4,082

Violent Crime is a rate (per 100,000) people. All other variables are in terms of share of the population (out of 100). N is the number of municipality-year observations. Violent crime data comes from Uniform Crime Reporting (UCR); all other data comes from Flood et al. (2020).

Table 6: Effects of Three Strikes Laws on Violent Crime Rates (During Matching, Control Treatment Year Set to Average Treatment Year in Full Sample)

	(1)	(2)	(3)	(4)
policy	20.3 [-28.7 , 71.5]	26.6 [-27.0 , 83.7]	31.7 [-10.1 , 86.1]	30.8 [-5.3 , 77.4]
Municipality FE	X	X	X	X
Border-Region x Timing FE	X	X	X	X
County Controls		X		X
Border-Region Weights	X	X		
Matching Weights			X	X
$N$	12,103	12,103	7,748	7,748
Mean Dependent Variable	428	428	371	371
No. of clusters	46	46	43	43
P-value ( $H_0 : \beta = 0$ )	0.439	0.384	0.144	0.105
P-value ( $H_0 : \beta < 0$ )	0.224	0.188	0.066*	0.051*

Notes: \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a three strikes law and is in terms of a rate per 100,000 people. Wild cluster bootstrap confidence intervals clustered at the state level are in square brackets and are calculated using 1,000 replications. County controls include male population share, black population share, Hispanic population share, share of population aged 18-44, share of population in poverty, share of population that is unemployed, and share of population without a high school degree.  $N$  is the number of municipality-year observations. Matching is done via 1:N “nearest-neighbor” propensity score matching (with replacement), enforcing common support, with five controls matched to each treatment, and a caliper set equal to 0.05.



Table 7: Effects of Three Strikes Laws on Violent Crime Rates (Excludes Municipalities from Border Regions Consisting of Two Treatment States with Heterogeneous Treatment Timings)

	(1)	(2)	(3)	(4)
policy	23.2 [-28.3 , 77.9]	31.6 [-24.4 , 93.3]	62.7 [-2.9 , 131.1]	65.6 [17.7 , 127.7]
Municipality FE	X	X	X	X
Border-Region x Year FE	X	X	X	X
County Controls		X		X
Border-Region Weights	X	X		
Matching Weights			X	X
$N$	10,829	10,829	8,151	8,151
Mean Dependent Variable	410	410	396	396
No. of clusters	44	44	41	41
P-value ( $H_0 : \beta = 0$ )	0.425	0.293	0.061*	0.005***
P-value ( $H_0 : \beta < 0$ )	0.207	0.153	0.030**	0.003***

Notes: \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a three strikes law and is in terms of a rate per 100,000 people. Wild cluster bootstrap confidence intervals clustered at the state level are in square brackets and are calculated using 1,000 replications. County controls include male population share, black population share, Hispanic population share, share of population aged 18-44, share of population in poverty, share of population that is unemployed, and share of population without a high school degree.  $N$  is the number of municipality-year observations. Matching is done via 1:N “nearest-neighbor” propensity score matching (with replacement), enforcing common support, with five controls matched to each treatment.

Table 8: Effects of Three Strikes Laws on Violent Crime Rates (Violent Crime Rate Includes Rapes)

	(1)	(2)	(3)	(4)
policy	23.6 [-27.0 , 78.5]	30.6 [-27.2 , 90.1]	43.2 [-22.8 , 107.7]	48.4 [-5.2 , 118.4]
Municipality FE	X	X	X	X
Border-Region x Year FE	X	X	X	X
County Controls		X		X
Border-Region Weights	X	X		
Matching Weights			X	X
$N$	12,103	12,103	7,995	7,995
Mean Dependent Variable	466	466	446	446
No. of clusters	46	46	42	42
P-value ( $H_0 : \beta = 0$ )	0.404	0.330	0.150	0.070*
P-value ( $H_0 : \beta < 0$ )	0.207	0.147	0.071*	0.025**

Notes: \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a three strikes law and is in terms of a rate per 100,000 people. Wild cluster bootstrap confidence intervals clustered at the state level are in square brackets and are calculated using 1,000 replications. County controls include male population share, black population share, Hispanic population share, share of population aged 18-44, share of population in poverty, share of population that is unemployed, and share of population without a high school degree.  $N$  is the number of municipality-year observations. Matching is done via 1:N “nearest-neighbor” propensity score matching (with replacement), enforcing common support, with five controls matched to each treatment.

Table 9: Effects of Three Strikes Laws on Violent Crime Rates (Sample Restricted to States who Frequently Applied Enhanced Sentencing)

	(1)	(2)	(3)	(4)
policy	-9.2 [-158.6 , 118.6]	-10.5 [-226.2 , 195.3]	12.3 [-313.7 , 167.1]	-12.3 [-142.0 , 191.0]
Municipality FE	X	X	X	X
Border-Region x Year FE	X	X	X	X
County Controls		X		X
Border-Region Weights	X	X		
Matching Weights			X	X
$N$	1,963	1,963	1,183	1,183
Mean Dependent Variable	717	717	673	673
No. of clusters	11	11	10	10
P-value ( $H_0 : \beta = 0$ )	0.898	0.892	0.939	0.815
P-value ( $H_0 : \beta < 0$ )	0.552	0.571	0.474	0.579

Notes: \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a three strikes law and is in terms of a rate per 100,000 people. Wild cluster bootstrap confidence intervals clustered at the state level are in square brackets and are calculated using 1,000 replications. County controls include male population share, black population share, Hispanic population share, share of population aged 18-44, share of population in poverty, share of population that is unemployed, and share of population without a high school degree.  $N$  is the number of municipality-year observations. Matching is done via 1:N “nearest-neighbor” propensity score matching (with replacement), enforcing common support, with five controls matched to each treatment.