

THREE ESSAYS ON CRIME POLICY
AND THE BAYESIAN BOOTSTRAP

A Dissertation
presented to
the Faculty of the Graduate School
at the University of Missouri

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
LONNIE HOFMANN
Dr. David Kaplan, Dissertation Supervisor

MAY 2021

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled

THREE ESSAYS ON CRIME POLICY
AND THE BAYESIAN BOOTSTRAP

presented by Lonnie Hofmann,

a candidate for the degree of Doctor of Philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor David Kaplan

Professor Cory Koedel

Professor Peter Mueser

Professor Dana Weston

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisors, Dr. David Kaplan and Dr. Cory Koedel. I would never have been able to finish this dissertation without their continuous help and invaluable input. I would also like to thank Dr. Peter Mueser for his helpful comments and advice, and Dr. Dana Weston for her support throughout my academic career at the University of Missouri.

Special thanks to Darlyne, Cliff, Gavin, and Randy for their continual contributions to my mental well-being.

Most of all, I would like to thank my parents, Steve and Wendy. Completion of this dissertation would have been impossible without their unconditional love.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	iv
LIST OF TABLES	v
ABSTRACT	vii
CHAPTER	
1 High-order Frequentist Coverage Probability of Bayesian Bootstrap Credible Intervals for Quantiles	1
2 The Effect of Three Strikes Laws on Violent Crime Rates	20
3 The Impact of Legalizing Medical Marijuana on Violent Crime Rates	50
APPENDIX	
A Appendix for Chapter 1	75
B Appendix for Chapter 2	78
C Appendix for Chapter 3	87
BIBLIOGRAPHY	89
VITA	97

LIST OF FIGURES

1.1	Justification for Equation (1.9)	13
2.1	States that Passed a Three Strikes Law and the Year of Passage	43
2.2	Border Municipality Pre-Trend Analysis, Three Strikes Laws	45
2.3	Interior Municipality Pre-Trend Analysis, Three Strikes Laws	47
3.1	States that Legalized Medical Marijuana from 1996–2003 and the Year of Implementation	69
3.2	Border Municipality Pre-Trend Analysis, Medical Marijuana	71
3.3	Interior Municipality Pre-Trend Analysis, Medical Marijuana	73

LIST OF TABLES

1.1	Simulated coverage probability (CP)	18
2.1	Summary Statistics (Full Sample), Three Strikes Laws	44
2.2	Effects of Three Strikes Laws on Violent Crime Rates, Border Municipality Analysis	46
2.3	Effects of Three Strikes Laws on Violent Crime Rates, Interior Municipality Analysis	48
2.4	Effects of Three Strikes Laws on Other Outcomes	49
3.1	Summary Statistics (After Matching), Medical Marijuana	70
3.2	Effects of Medical Marijuana Legalization on Adjusted Violent Crime Rates, Border Municipality Analysis	72
3.3	Effects of Medical Marijuana Legalization on Adjusted Violent Crime Rates, Interior Municipality Analysis	74
B.1	Summary Statistics (After Matching), Three Strikes Laws	80
B.2	Effects of Three Strikes Laws on Violent Crime Rates (Sample Includes All Border Municipalities)	81
B.3	Effects of Three Strikes Laws on Violent Crime Rates (Confidence Intervals are Clustered at the Border-Region Level)	82
B.4	Effects of Three Strikes Laws on Violent Crime Rates (During Matching, Control Treatment Year Set to Average Treatment Year in Full Sample)	83

B.5	Effects of Three Strikes Laws on Violent Crime Rates (Excludes Municipalities from Border Regions Consisting of Two Treatment States with Heterogeneous Treatment Timings)	84
B.6	Effects of Three Strikes Laws on Violent Crime Rates (Violent Crime Rate Includes Rapes)	85
B.7	Effects of Three Strikes Laws on Violent Crime Rates (Sample Restricted to States who Frequently Applied Enhanced Sentencing) . . .	86
C.1	Summary Statistics (Full Sample), Medical Marijuana	88

THREE ESSAYS ON CRIME POLICY AND THE BAYESIAN BOOTSTRAP

Lonnie Hofmann

Dr. David Kaplan, Dissertation Supervisor

ABSTRACT

This dissertation consists of three chapters. In the first chapter, I analyze credible intervals for quantiles constructed using Bayesian bootstrap techniques and show that credible intervals constructed using the “continuity-corrected” Bayesian bootstrap (Banks, 1988) have frequentist coverage probability error of only $O(n^{-1})$. In addition, I show that these “continuity-corrected” Bayesian bootstrap credible intervals achieve the same frequentist coverage probability as the frequentist confidence intervals of Goldman and Kaplan (2017), up to some error term of magnitude $O(n^{-1})$. Furthermore, I demonstrate that credible intervals constructed using the “continuity-corrected” Bayesian bootstrap have less frequentist coverage probability error than those constructed using the Bayesian bootstrap (Rubin, 1981).

In the second chapter, I investigate three strikes laws, which mandate sharply increased sentences for criminals who commit a specific number of felonies. Specifically, I analyze the effect of these laws on violent crime rates using municipal-level data from the FBI. 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 no statistical evidence that three strikes laws reduce violent crime rates. I rule out reductions in violent crime rates greater than 1.3% and 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.

In the third chapter, I examine medical marijuana laws (MMLs), which legalize the use, possession, and cultivation of marijuana by individuals with qualifying medical conditions. Namely, I employ municipal-level data from the FBI to analyze the effect of MMLs on violent crime rates. I compare municipalities in border regions with different treatments statuses using a difference-in-differences specification with a sample matched on pre-treatment outcomes. I find a lack of evidence for MMLs increasing violent crime rates, but I cannot eliminate the possibility of small-to-medium positive effects. However, I rule out increases in violent crime rates greater than 9.9% and reject the hypothesis that MMLs increase violent crime at the 10% significance level.

Chapter 1

High-order Frequentist Coverage Probability of Bayesian Bootstrap Credible Intervals for Quantiles

1.1 Introduction

In this paper, I study credible intervals for quantiles constructed using Bayesian bootstrap techniques. The Bayesian bootstrap (BB) (Rubin, 1981) is a general, non-parametric, Bayesian method for approximating posterior distributions of population objects like quantiles. The BB only produces discrete posterior distributions, so it may be helpful to smooth the distributions with the “continuity-corrected” Bayesian bootstrap (CCBB) (Banks, 1988).

The main theorem in my paper shows that the CCBB has very accurate coverage. Constructing credible intervals for quantiles using the CCBB leads to frequentist coverage probability error (CPE) of only $O(n^{-1})$. This accuracy is better than

many frequentist methods, with exceptions being some methods, such as Goldman and Kaplan (2017), specific to iid data, and some bootstrap methods. Some bootstrap examples include Polansky and Schucany (1997), who achieve two-sided CPE of $O(n^{-3/2})$ but admit the shortcomings of their bandwidth estimation technique, and Ho and Lee (2005), among others. However, most of these methods are more computationally burdensome and are sensitive to bandwidth choices, and, even then, many of the methods do not achieve the same coverage accuracy as the CCBB. Thus, as the CCBB does not require computation of bandwidths, ease of computation along with very accurate coverage are two advantages to using the CCBB.

The CCBB results in my paper are closely related to the frequentist results from Goldman and Kaplan (2017), and I show that the frequentist coverage probability (CP) of credible intervals constructed using the CCBB is equal to the CP of the confidence intervals constructed in Goldman and Kaplan (2017) up to some error term of magnitude $O(n^{-1})$. This theoretical result is supported via simulations and implies that a confidence interval from Goldman and Kaplan (2017) can be interpreted as an approximate Bayesian credible interval.

I also show that credible intervals for quantiles constructed using the CCBB have less frequentist coverage probability error (CPE) than those constructed using the BB. Specifically, I show that the CCBB has CPE of magnitude $O(n^{-1})$ while the BB has CPE of magnitude $O(n^{-1/2})$. I confirm the CCBB's theoretical advantage via simulations, which demonstrate the increased coverage accuracy of using the CCBB instead of the BB.

In the literature on the Bayesian bootstrap, Rubin (1981) introduces the technique, and Banks (1988) extends Rubin's results to a continuous setting. Meeden (1993) provides another variation on Rubin's original Bayesian bootstrap via a grid-

based procedure. This grid-based procedure is a “pure” Bayesian method, whereas the CCBB is not. However, the grid-based procedure and the CCBB coincide if the grid coincides with the order statistics, so it may still be reasonable to interpret the CCBB as producing Bayesian posteriors. Meeden (1993) also provides simulations that offer evidence of the benefit to using the CCBB, but he does not provide a theoretical justification for why this may be. I use both the CCBB (Banks, 1988) and the BB (Rubin, 1981) in my analysis (but not the grid-based Bayesian bootstrap (Meeden, 1993)), and I provide some theoretical justification of the benefit to using the CCBB.

In the literature on quantile confidence intervals, Goldman and Kaplan (2017) construct confidence intervals for quantiles using order statistic theory and the probability integral transform of Fisher (1932), Pearson (1933), and Neyman (1937). Goldman and Kaplan (2017) build off of the work done by Hutson (1999), who had proposed using L-statistics, specifically linearly interpolated order statistics, to construct confidence intervals for quantiles. Various other methods have been used to construct such confidence intervals including Kaplan (2015), who uses Edgeworth expansions, and Beran and Hall (1993), who also use L-statistics but with an interpolation weight based on the binomial distribution. Goldman and Kaplan (2018) extend the results of Goldman and Kaplan (2017) to interquartile ranges, among other things.

This paper shows that high-order coverage accuracy of quantiles can be achieved through Bayesian bootstrap techniques and not just via frequentist techniques such as those used in Goldman and Kaplan (2017). In addition, since the confidence intervals in Goldman and Kaplan (2017) can be interpreted as approximate Bayesian credible intervals, this raises the possibility that high-order results could be achieved in settings other than the ones discussed in this paper. For example, extending the CCBB to settings involving clustering or survey data with complex sampling designs are

possibilities. Additionally, Hahn (1997) wondered about the ability of the Bayesian bootstrap to accurately approximate the distribution of the quantile regression estimator. Furthermore, Chamberlain and Imbens (2003) provide some motivation and examples for using Bayesian bootstrap methods in economic applications, including quantile regression.

Section 1.2 contains the theoretical results for single quantile inference. Section 1.3 contains simulation results. Section 1.4 suggests directions for future research.

Notation For scalar/vector variable formatting, \mathbf{X} is a random vector with elements X_j , and Y and y are random and non-random scalars, respectively. Random variables and vectors are uppercase (Y , X , etc.), while non-random values are lowercase (y , x). Also, $Q_\tau(\cdot)$ is the τ -quantile, $P(\cdot)$ probability, and \doteq “is equal to, up to smaller-order terms”. Acronyms used include those for Bayesian bootstrap (BB), continuity-corrected Bayesian bootstrap (CCBB), coverage probability (CP), coverage probability error (CPE), cumulative distribution function (CDF), and probability density function (PDF).

1.2 Main results

In this section, I introduce the “continuity-corrected” Bayesian bootstrap (CCBB) (Banks, 1988) along with the Bayesian bootstrap (BB) (Rubin, 1981) and present my main results for single quantile inference using the aforementioned Bayesian bootstrap techniques.¹

¹Banks refers to the CCBB as the “histospline-smoothed” Bayesian bootstrap in his paper.

1.2.1 Setup, notation, and intuition

BB and CCBB credible intervals for quantiles are constructed as follows. Consider an iid sample, $\mathbf{Y} = (Y_1, \dots, Y_n)'$, remembering that the data is fixed in a Bayesian framework. The BB uses an uninformative, improper Dirichlet process prior. Once we view \mathbf{Y} , we update the Dirichlet process prior to give us the posterior distribution of the *probabilities* of drawing specific elements from \mathbf{Y} . So, in the BB case, the posterior distribution of the population probability mass function is

$$p_1, \dots, p_n \mid \mathbf{Y} \sim \text{Dir}(1, \dots, 1), \quad (1.1)$$

where p_k is the probability of obtaining $Y_{n:k}$, the k th order statistic (i.e., k th smallest sample value).

The CCBB “smooths” the probabilities to give a continuous version of the BB posterior distributions. That is, in the CCBB case, the posterior distribution is

$$p_1, \dots, p_{n+1} \mid \mathbf{Y} \sim \text{Dir}(1, \dots, 1), \quad (1.2)$$

where p_k is the probability of obtaining any number between $Y_{n:k-1}$ and $Y_{n:k}$.

Since the probabilities associated with our sample have flat Dirichlet posteriors, we can use a result from Wilks (1962) to obtain distributions for various sums of these probabilities.² Given (1.1), Wilks (1962) shows that, for any k ,

$$p_1 + \dots + p_k \mid \mathbf{Y} \sim \text{Beta}(k, n - k). \quad (1.3)$$

²A flat Dirichlet distribution is defined as a Dirichlet distribution with all of its concentration parameters equal to one.

Additionally, given (1.2), the n in (1.3) changes to a $n + 1$ so that

$$p_1 + \cdots + p_k \mid \mathbf{Y} \sim \text{Beta}(k, n + 1 - k). \quad (1.4)$$

Now, consider the BB situation, i.e., (1.1), where p_k is the probability of drawing $Y_{n:k}$ from population CDF $F_Y(\cdot)$. Let $\pi_u(\cdot \mid \mathbf{Y})$ be the unsmoothed BB posterior probability conditional on \mathbf{Y} . I am interested in computing

$$\pi_u(Y_{n:k} < Q_\tau(Y) \mid \mathbf{Y}), \quad (1.5)$$

where $Q_\tau(Y) = \inf\{y : F_Y(y) \geq \tau\}$, i.e., the posterior probability or credibility of a one-sided, lower-end-point credible interval for the quantile of interest.³

To provide intuition for calculating (1.5), consider the following specific example where $n = 4$, and the quantile index $\tau = 0.5$. Since I am interested in computing $\pi_u(Y_{4:k} < Q_{0.5}(Y) \mid \mathbf{Y})$, I want to determine when $Y_{4:k} < Q_{0.5}(Y)$. First, it will be convenient to consider when $Y_{4:k} = Q_{0.5}(Y)$ for $k = 1, \dots, 4$. Since the BB setting is discrete, $Q_{0.5}(Y)$ can only take four values: $Y_{4:1}$, $Y_{4:2}$, $Y_{4:3}$, or $Y_{4:4}$.

First, $Q_{0.5}(Y) = Y_{4:1}$ if and only if $p_1 \geq 0.5$. In other words, in order for the median to equal the first order statistic, the probability of drawing the first order statistic, p_1 , must be greater than 0.5. In the extreme case where $p_1 = 1$, for example, the only value drawn will be $Y_{4:1}$, and $Q_{0.5}(Y) = Y_{4:1}$.

Furthermore, $Q_{0.5}(Y) = Y_{4:2}$ if and only if $p_1 + p_2 \geq 0.5$ and $p_1 < 0.5$. In other words, in order for the median to equal the second order statistic, the probability of drawing $Y_{4:2}$, or something less than $Y_{4:2}$ (here this means $Y_{4:1}$), must be greater than 0.5; but, additionally, the probability of drawing $Y_{4:1}$ must be less than 0.5 (otherwise

³The results presented in this paper are for one-sided, lower-end-point credible intervals but can easily be extended to one-sided, upper-end-point and two-sided credible intervals as well.

$Y_{4:1}$ would be the median).

Similarly, I can conclude that $Q_{0.5}(Y) = Y_{4:3}$ if and only if $p_1 + p_2 + p_3 \geq 0.5$ and $p_1 + p_2 < 0.5$, and $Q_{0.5}(Y) = Y_{4:4}$ if and only if $p_1 + p_2 + p_3 < 0.5$.⁴

Now consider the inequality $Y_{4:2} < Q_{0.5}(Y)$. Since $Q_{0.5}(Y) = Y_{4:2}$ if and only if $p_1 + p_2 \geq 0.5$ and $p_1 < 0.5$, then $Q_{0.5}(Y) \neq Y_{4:2}$ if and only if $p_1 + p_2 < 0.5$ or $p_1 \geq 0.5$. In the case where $p_1 \geq 0.5$, $Y_{4:1}$ is the median, so, $Y_{4:2} > Q_{0.5}(Y)$. Thus,

$$Y_{4:2} < Q_{0.5}(Y) \iff p_1 + p_2 < 0.5.$$

Similarly,

$$Y_{4:1} < Q_{0.5}(Y) \iff p_1 < 0.5,$$

$$Y_{4:3} < Q_{0.5}(Y) \iff p_1 + p_2 + p_3 < 0.5, \text{ and}$$

$$Y_{4:4} < Q_{0.5}(Y) \iff p_1 + p_2 + p_3 + p_4 < 0.5.$$

In other words,

$$Y_{4:k} < Q_{0.5}(Y) \iff \sum_{i=1}^k p_i < 0.5 \tag{1.6}$$

for $k = 1, \dots, 4$. Equation (1.6) is extrapolated to any n and any quantile index, τ , in Theorem 1.

1.2.2 Assumptions and theorems

I now list the main assumptions used in my paper. Assumption 1 lets me use (1.1) and (1.2), i.e., Dirichlet posterior distributions that treat all observations equally. Assumption 2 ensures that implementing the CCBB results in a posterior that generates only continuous CDFs.

⁴The condition $p_1 + p_2 + p_3 + p_4 \geq 0.5$ is always true since it is equal to one by definition.

Assumption 1. Sampling is iid: $Y_i \stackrel{iid}{\sim} F_Y(\cdot)$.

Assumption 2. Sample values are unique.

I now state Theorem 1 pointing out that Efron (1982) provides a similar result to my Theorem 1 in his book.

Theorem 1. *Under Assumptions 1 and 2 and (1.1), and for $k = 1, \dots, n$,*

$$\pi_u(Y_{n:k} < Q_\tau(Y) \mid \mathbf{Y}) = F_\beta^1(\tau),$$

where F_β^1 is the CDF of $\text{Beta}(k, n - k)$.

Proof. First,

$$Y_{n:k} < Q_\tau(Y) \iff \sum_{i=1}^k p_i < \tau.$$

Thus,

$$\pi_u(Y_{n:k} < Q_\tau(Y) \mid \mathbf{Y}) = \pi_u\left(\sum_{i=1}^k p_i < \tau \mid \mathbf{Y}\right).$$

By (1.3), $\sum_{i=1}^k p_i \mid \mathbf{Y} \sim \text{Beta}(k, n - k)$. Thus, $\pi_u(\sum_{i=1}^k p_i < \tau \mid \mathbf{Y}) = F_\beta^1(\tau)$, by properties of a continuous CDF. \square

Theorem 1 is useful because, in order to compute the credibility of a one-sided, lower-end-point credible interval for the quantile of interest, I only need to evaluate a specific beta CDF. Evaluating such a beta CDF can be done easily in any common statistical software.

In addition to computing $\pi_u(Y_{n:k} < Q_\tau(Y) \mid \mathbf{Y})$, as presented in Theorem 1, I am also interested in computing

$$\pi_u(Y_{n:k} \leq Q_\tau(Y) \mid \mathbf{Y}) \tag{1.7}$$

since the BB is discrete.

For intuition in computing (1.7), consider once again the example where $n = 4$ and $\tau = 0.5$. From the discussion in Section 1.2.1, recall that $Y_{4:2} > Q_{0.5}(Y)$ when $p_1 \geq 0.5$ which implies

$$Y_{4:2} \leq Q_{0.5}(Y) \iff p_1 < 0.5.$$

To determine when $Y_{4:3} \leq Q_{0.5}(Y)$, recall from the discussion in Section 1.2.1 that $Q_{0.5}(Y) = Y_{4:3}$ if and only if $p_1 + p_2 + p_3 \geq 0.5$ and $p_1 + p_2 < 0.5$. So, $Q_{0.5}(Y) \neq Y_{4:3}$ if and only if $p_1 + p_2 + p_3 < 0.5$ or $p_1 + p_2 \geq 0.5$. The proof of Theorem 1 implies $Y_{4:3} < Q_{0.5}(Y)$ if and only if $p_1 + p_2 + p_3 < 0.5$, so $Y_{4:3} > Q_{0.5}(Y)$ if and only if $p_1 + p_2 \geq 0.5$. Thus,

$$Y_{4:3} \leq Q_{0.5}(Y) \iff p_1 + p_2 < 0.5.$$

Similarly,

$$Y_{4:4} \leq Q_{0.5}(Y) \iff p_1 + p_2 + p_3 < 0.5.$$

In other words,

$$Y_{4:k} \leq Q_{0.5}(Y) \iff \sum_{i=1}^{k-1} p_i < 0.5 \tag{1.8}$$

for $k = 1, \dots, 4$. Equation (1.8) is extrapolated to any n and any quantile index, τ , in Theorem 2.

Theorem 2. *Under Assumptions 1 and 2 and (1.1), and for $k = 1, \dots, n$,*

$$\pi_u(Y_{n:k} \leq Q_\tau(Y) \mid \mathbf{Y}) = F_\beta^2(\tau),$$

where F_β^2 is the CDF of $\text{Beta}(k-1, n+1-k)$.

Proof. First,

$$Y_{n:k} \leq Q_\tau(Y) \iff \sum_{i=1}^{k-1} p_i < \tau.$$

Thus,

$$\pi_u(Y_{n:k} \leq Q_\tau(Y) \mid \mathbf{Y}) = \pi_u\left(\sum_{i=1}^{k-1} p_i < \tau \mid \mathbf{Y}\right).$$

By (1.3), $\sum_{i=1}^{k-1} p_i \mid \mathbf{Y} \sim \text{Beta}(k-1, n+1-k)$. Thus, $\pi_u(\sum_{i=1}^{k-1} p_i < \tau \mid \mathbf{Y}) = F_\beta^2(\tau)$, by properties of a continuous CDF. \square

In Theorem 1 and Theorem 2, I have presented results for computing the posterior probability of one-sided, lower-end-point credible intervals for the quantile of interest in the BB setting. In Theorem 3, we extend these results to the CCBB setting. Given an iid sample $\mathbf{Y} = (Y_1, \dots, Y_n)'$, let $\pi_s(\cdot \mid \mathbf{Y})$ be the CCBB posterior probability conditional on \mathbf{Y} , where s stands for smoothed.

Theorem 3. *Under Assumptions 1 and 2 and (1.2), and for $k = 1, \dots, n$,*

$$\pi_s(Y_{n:k} < Q_\tau(Y) \mid \mathbf{Y}) = \pi_s(Y_{n:k} \leq Q_\tau(Y) \mid \mathbf{Y}) = F_\beta^3(\tau),$$

where F_β^3 is the CDF of $\text{Beta}(k, n+1-k)$.

Proof. First,

$$Y_{n:k} < Q_\tau(Y) \iff \sum_{i=1}^k p_i < \tau.$$

So,

$$\pi_s(Y_{n:k} < Q_\tau(Y) \mid \mathbf{Y}) = \pi_s\left(\sum_{i=1}^k p_i < \tau \mid \mathbf{Y}\right).$$

Also,

$$Y_{n:k} \leq Q_\tau(Y) \iff \sum_{i=1}^k p_i \leq \tau.$$

So,

$$\pi_s(Y_{n:k} \leq Q_\tau(Y) \mid \mathbf{Y}) = \pi_s\left(\sum_{i=1}^k p_i \leq \tau \mid \mathbf{Y}\right).$$

Thus,

$$\pi_s(Y_{n:k} < Q_\tau(Y) \mid \mathbf{Y}) = \pi_s\left(\sum_{i=1}^k p_i < \tau \mid \mathbf{Y}\right) = \pi_s\left(\sum_{i=1}^k p_i \leq \tau \mid \mathbf{Y}\right) = \pi_s(Y_{n:k} \leq Q_\tau(Y) \mid \mathbf{Y}).$$

By (1.4), $\sum_{i=1}^k p_i \mid \mathbf{Y} \sim \text{Beta}(k, n + 1 - k)$. Thus, $\pi_s\left(\sum_{i=1}^k p_i < \tau \mid \mathbf{Y}\right) = F_\beta^3(\tau)$, by properties of a continuous CDF. \square

The posterior probability that the τ -quantile lies in the interval $[Y_{n:k}, \infty)$ (as seen in Theorem 3) is identical to the frequentist coverage probability of the same interval for the same quantile. That is, any interval whose endpoint is an order statistic has exactly the same credibility level and confidence level.⁵

Up to this point, Theorems 1–3 have all used order statistics as endpoints for the one-sided, lower-end-point credible intervals I have considered. However, none of these order statistics have been *fractional*. That is, we have not considered the case when k is not an integer. Considering fractional order statistics will be useful as sometimes the credible interval for the credibility level of interest (e.g., 95%) requires a fractional order statistic as an endpoint. Theorem 4 extends the results of Theorem 3 to such fractional order statistics.

Prior to presenting Theorem 4, I present some fractional order statistic notation. For $u \in (0, 1)$, let $k \equiv \lfloor u(n + 1) \rfloor$ and $\epsilon \equiv u(n + 1) - k$, where $\lfloor \cdot \rfloor$ is the floor function and $\epsilon \in [0, 1)$ is the interpolation weight. Thus, $Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k})$ is the fractional order statistic between $Y_{n:k}$ and $Y_{n:k+1}$ with interpolation weight ϵ .

Additionally, to aid my proof of Theorem 4, I write the quantile corresponding to

⁵This property also holds for two-sided intervals where both endpoints are order statistics.

a posterior draw of (p_1, \dots, p_{n+1}) as:

$$Q_\tau(Y) = \frac{(\tau - \sum_{i=1}^k p_i)(Y_{n:k+1} - Y_{n:k})}{p_{k+1}} + Y_{n:k}, \quad (1.9)$$

where k equals the maximum integer such that $\sum_{i=1}^k p_i \leq \tau$. Justification for Equation (1.9) can be seen in Figure 1.1 which is underpinned by the piecewise-linear smoothing used by Banks (1988) for the CCBB.

Figure 1.1 considers the specific situation where $n = 2$ and $k = 1$. In the figure, the solid, black, connected line segments represent the CDF generated by a posterior draw from (1.2). In particular, the diagonal, solid, black line between $Y_{n:1}$ and $Y_{n:2}$ is a result of applying the CCBB to a BB posterior distribution by spreading the probability, p_2 , throughout the interval $(Y_{n:1}, Y_{n:2}]$. In addition, the dotted, vertical line at Q_τ represents the quantile of interest. Finally, the dotted, horizontal line at p_1 represents the probability of obtaining any number in $(-\infty, Y_{n:1}]$. Taken together, the three aforementioned lines form a right triangle. As such, I make use of the fact that the slope of a line is equal to its “rise” over its “run” (in this case, the line we are considering is the diagonal, solid black line between $Y_{n:1}$ and $Y_{n:2}$).

Let slope = s , rise = R , and run = r . In Figure 1.1, $R = \tau - p_1$ and $s = p_2/(Y_{n:2} - Y_{n:1})$. Furthermore, $r = Q_\tau(Y) - Y_{n:1}$, so $Q_\tau(Y) = Y_{n:1} + r$. Finally, $r = R/s = (\tau - p_1)(Y_{n:2} - Y_{n:1})/p_2$ which implies that

$$Q_\tau(Y) = \frac{(\tau - p_1)(Y_{n:2} - Y_{n:1})}{p_2} + Y_{n:1}. \quad (1.10)$$

We can extrapolate (1.10) to any n and any k in order to get (1.9).

In addition to (1.9), I present (as Assumption 3) “Assumption A2” from Goldman and Kaplan (2017), almost verbatim, prior to presenting Theorem 4. Assumption 3 is necessary to ensure the population distribution is well-behaved enough for Goldman

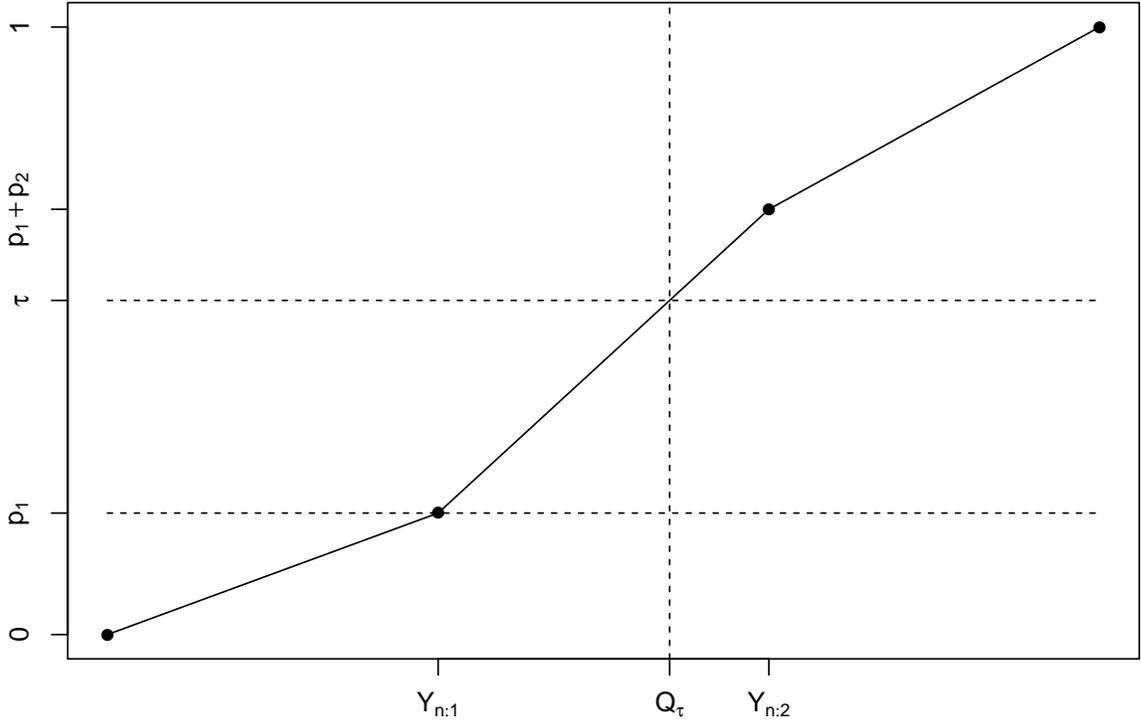


Figure 1.1: Justification for Equation (1.9)

and Kaplan (2017) to obtain their results. In addition, Assumption 3 is necessary in my paper as some results from Goldman and Kaplan (2017) are used in the proof of Theorem 4.

Assumption 3. For each quantile τ , the PDF $f_y(\cdot)$ (corresponding to CDF $F_y(\cdot)$ in Assumption 1) satisfies:

- (i) $f_y(F_y^{-1}(\tau)) > 0$;
- (ii) $f_y''(\cdot)$ is continuous in some neighborhood of $F_y^{-1}(\tau)$, i.e., $f_y \in C^2(U_\delta(F_y^{-1}(\tau)))$ with $U_\delta(x)$ denoting some δ -neighborhood of point $x \in \mathbb{R}$.

Theorem 4. Let $k \equiv \lfloor u(n+1) \rfloor$ and $\epsilon \equiv u(n+1) - k$, where $\lfloor \cdot \rfloor$ is the floor function and ϵ is the interpolation weight. Under Assumptions 1–3 and (1.2), and for any

$$u = \tau + o(1),$$

$$\pi_s[Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k}) < Q_\tau(Y) \mid \mathbf{Y}] = F_\beta^4(\tau) + O(n^{-1}),$$

where F_β^4 is the CDF of $\text{Beta}(k + \epsilon, n + 1 - k - \epsilon)$.

Proof. By (1.9), $Q_\tau(Y) = (\tau - \sum_{i=1}^c p_i)(Y_{n:c+1} - Y_{n:c})/p_{c+1} + Y_{n:c}$, where c equals the maximum integer such that $\sum_{i=1}^c p_i \leq \tau$. Thus, in the situation where $c = k$,

$$\begin{aligned} & \pi_s[Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k}) < Q_\tau(Y) \mid \mathbf{Y}] \\ &= \pi_s[Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k}) < (\tau - \sum_{i=1}^k p_i)(Y_{n:k+1} - Y_{n:k})/p_{k+1} + Y_{n:k} \mid \mathbf{Y}] \\ &= \pi_s[\epsilon(Y_{n:k+1} - Y_{n:k}) < (\tau - \sum_{i=1}^k p_i)(Y_{n:k+1} - Y_{n:k})/p_{k+1} \mid \mathbf{Y}] \\ &= \pi_s(\epsilon p_{k+1} < \tau - \sum_{i=1}^k p_i \mid \mathbf{Y}) \\ &= \pi_s(\sum_{i=1}^k p_i + \epsilon p_{k+1} < \tau \mid \mathbf{Y}). \end{aligned}$$

In other words,

$$\pi_s[Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k}) < Q_\tau(Y) \mid \mathbf{Y}] \iff \pi_s(\sum_{i=1}^k p_i + \epsilon p_{k+1} < \tau \mid \mathbf{Y}). \quad (1.11)$$

When $c < k$, (1.11) holds because $Q_\tau(Y)$ is never larger than $Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k})$, i.e., $Q_\tau(Y) < Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k})$ always. In addition, when $c > k$, (1.11) holds because $Q_\tau(Y)$ is always larger than $Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k})$, i.e., $Q_\tau(Y) > Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k})$ always.

Now, suppose $U_i \stackrel{iid}{\sim} \text{Unif}(0, 1)$ for $i = 1, \dots, n$, and let $U_{n:0} \equiv 0$ and $U_{n:n+1} \equiv 1$.

Then,

$$(U_{n:1} - U_{n:0}, U_{n:2} - U_{n:1}, \dots, U_{n:n} - U_{n:n-1}, U_{n:n+1} - U_{n:n}) \sim \text{Dir}(1, \dots, 1), \quad (1.12)$$

where the above sampling Dirichlet distribution has $n + 1$ parameters (Wilks, 1962).

I then apply Theorem 2 from Goldman and Kaplan (2017) to deduce that

$$\text{P}[U_{n:k} + \epsilon(U_{n:k+1} - U_{n:k}) < \tau] = F_{\beta}^4(\tau) + O(n^{-1}), \quad (1.13)$$

where $F_{\beta}^4(\tau)$ is the CDF of $\text{Beta}(k + \epsilon, n + 1 - k - \epsilon)$ as presented in the statement of the theorem.

Recall that $p_1, \dots, p_{n+1} \mid \mathbf{Y} \sim \text{Dir}(1, \dots, 1)$, by (1.2), which is identical to the joint distribution of the $U_{n:k+1} - U_{n:k}$ in (1.12). This implies that

$$\pi_s\left(\sum_{i=1}^k p_i + \epsilon p_{k+1} < \tau \mid \mathbf{Y}\right) = F_{\beta}^4(\tau) + O(n^{-1}), \quad (1.14)$$

where I switch out P for $\pi_s(\cdot \mid \mathbf{Y})$ since the interpretation changes from a sampling probability to a posterior probability.

Finally, (1.14) implies that

$$\pi_s[Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k}) < Q_{\tau}(Y) \mid \mathbf{Y}] = F_{\beta}^4(\tau) + O(n^{-1}). \quad \square$$

The implication from Theorem 4, coupled with Theorem 4 from Goldman and Kaplan (2017), is that the coverage probability (CP) of a one-sided, lower-end-point *confidence* interval for the quantile of interest and the conditional posterior probability of the one-sided, lower-end-point *credible* interval for the quantile of interest are equal up to $O(n^{-1})$. Specifically,

$$\text{P}[Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k}) < Q_{\tau}(Y)] + O(n^{-1})$$

$$\begin{aligned}
&= \pi_s[Y_{n:k} + \epsilon(Y_{n:k+1} - Y_{n:k}) < Q_\tau(Y) \mid \mathbf{Y}] + O(n^{-1}) \\
&= F_\beta^4(\tau).
\end{aligned}$$

As Theorem 4 uses the CCBB to form credible intervals, and the CCBB is more complicated than the BB, it is necessary to show that there is some benefit in using the CCBB to construct credible intervals instead of the BB. I show in Theorem 5 that using the CCBB to construct credible intervals, instead of the BB, results in credible intervals with smaller CPE.

Prior to presenting Theorem 5, I introduce some additional notation. First let $1 - \alpha = F_\beta^4(\tau)$, where $F_\beta^4(\tau)$ is defined in Theorem 4. Also, define $u^l(\alpha) \equiv (k + \epsilon)/(n + 1)$. Thus, $u^l(\alpha)$ encapsulates the k , ϵ , and n necessary so that $1 - \alpha = F_\beta^4(\tau)$. Additionally, let $z_{1-\alpha}$ denote the $(1 - \alpha)$ -quantile of a standard normal distribution, i.e., $z_{1-\alpha} \equiv \Phi^{-1}(1 - \alpha)$.

Lemma 3 from Goldman and Kaplan (2017) allows me to approximate $u^l(\alpha)$ using quantiles of the standard normal distribution as follows:

$$u^l(\alpha) = \tau - n^{-1/2} z_{1-\alpha} \sqrt{\tau(1-\tau)} - \frac{(2\tau-1)}{6n} (z_{1-\alpha}^2 + 2) + O(n^{-3/2}).$$

The idea behind Theorem 5 is that I want to show that there is some number c such that:

$$1 - \alpha + \frac{c}{\sqrt{n}} = \pi_s(\text{Beta}(k, n + 1 - k) < \tau) + O(n^{-1}).^6$$

In other words, changing $u^l(\alpha)$ from $\frac{k+\epsilon}{n+1}$ to $\frac{k}{n+1}$, i.e., changing $u^l(\alpha)$ by $\frac{\epsilon}{n+1}$ (or essentially $\frac{\epsilon}{n}$ as n gets large) results in the CP changing by $\frac{c}{\sqrt{n}}$. That is, using the BB instead of the CCBB to construct credible intervals for the quantile of interest

⁶Note, via Theorem 1, that the BB actually uses $\text{Beta}(k, n - k)$ as opposed to $\text{Beta}(k, n + 1 - k)$ but this is just a difference of $O(n^{-1})$, which doesn't matter in this context.

results in CPE of order of magnitude $O(n^{-1/2})$. I now present Theorem 5.

Theorem 5. *Using the BB to construct credible intervals instead of the CCBB results in CPE of $O(n^{-1/2})$.*

Proof. Compiled in Appendix A.1. □

The implication from Theorem 5 is that there is an increase in CPE (by c/\sqrt{n}) when using the BB to construct credible intervals instead of the CCBB, i.e., changing $u^l(\alpha)$ by $\frac{c\sqrt{\tau(1-\tau)}}{n\phi(z_{1-\alpha})}$ (i.e., something over n).

1.3 Simulation study

In this section, I present results from a simulation study that compares the coverage probability (CP) of three different types of two-sided intervals: credible intervals constructed using both the “continuity-corrected” Bayesian bootstrap (CCBB) and the Bayesian bootstrap (BB), and confidence intervals constructed using the “L-statistic” method (GK).⁷

The simulations were done in R (R Core Team, 2017) using a sample size of $n = 25$, 10,000 posterior draws from the Dirichlet posterior (for the CCBB and the BB), and 10,000 simulation replications from, F , the relevant sampling distribution (for all three methods). The DGP is taken from Table 2 of Goldman and Kaplan (2017). The CCBB and BB methods use code provided by Martin, Quinn, and Park (2011) for draws from the Dirichlet posterior distribution.

Table 1.1 shows that the CCBB and GK perform similarly when $n = 25$ and $\tau = 0.5$ in that they achieve CP close to the 95 percent target level, whereas the BB

⁷This “L-statistic” method was originally proposed by Hutson (1999) and examined further by Goldman and Kaplan (2017).

Table 1.1: Simulated coverage probability (CP)

n	τ	F	CP		
			BB	CCBB	GK
25	0.5	Normal	0.9415	0.9531	0.9547
25	0.5	Uniform	0.9388	0.9472	0.9489
25	0.5	Exponential	0.9412	0.9475	0.9502

$1 - \alpha = 0.95$; n , τ , and distributions of $Y_i(F)$ shown in table; 10,000 replications. Normal, Uniform, and Exponential refer to the standard normal distribution, standard uniform distribution, and standard exponential distribution, respectively.

consistently undercovers by a relatively larger amount. GK tends to undercover less than the CCBB because the $O(n^{-1})$ error term associated with GK is always positive (Goldman and Kaplan, 2017), whereas the $O(n^{-1})$ error term associated with the CCBB could be positive or negative.⁸

1.4 Conclusion and future work

I analyze credible intervals for quantiles constructed using Bayesian bootstrap techniques. I show that the “continuity-corrected” Bayesian bootstrap (CCBB) (Banks, 1988) has very accurate coverage, in that constructing credible intervals for quantiles using the CCBB leads to frequentist coverage probability error (CPE) of only $O(n^{-1})$. Simulation results reflect the high-order coverage accuracy of the smoothed Bayesian bootstrap credible intervals.

In addition, I prove that the frequentist coverage probability (CP) of the CCBB credible intervals and the CP of the confidence intervals constructed in Goldman and Kaplan (2017) are the same up to some error term of magnitude $O(n^{-1})$. This finding

⁸Using GK or CCBB to construct confidence or credible intervals, respectively, results in an $O(n^{-1})$ error term, in terms of CP, which is what is referenced above.

is supported via simulations.

I also prove that constructing credible intervals using the CCBB in place of the Bayesian bootstrap (BB) (Rubin, 1981) leads to credible intervals with frequentist coverage probability error (CPE) of magnitude $O(n^{-1})$ as opposed to magnitude $O(n^{-1/2})$. This conclusion is also supported via simulations. Additional simulations could be done that incorporate the grid-based variation of the Bayesian bootstrap proposed by Meeden (1993).

Possible future work could include relaxing the iid assumption by allowing for sampling weights using the weighted gamma process prior of Lo (1993). The hope is that desired frequentist properties can be achieved even when the iid assumption is relaxed. Other work could involve extending the results of this paper to clustering, conditional quantiles, or quantile differences/interquartile ranges. It would be of interest to see if using the CCBB leads to high-order accuracy for quantile models beyond the special case outlined in this paper.

Chapter 2

The Effect of Three Strikes Laws on Violent Crime Rates

2.1 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. They also aim to 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.

This paper analyzes the effect of three strikes laws on violent crime rates using municipal-level data from the FBI. 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 no statistical evidence that three strikes laws 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.

My paper makes two contributions to the literature on three strikes laws. First, it analyzes the impact of three strikes laws on crime outcomes using municipal level data in all states that passed such a law. Many studies have analyzed the effects of a three strikes law passed in California in 1994, as California's law is 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 reduced crime. For example, Datta (2017) and Shepherd (2002) use instrumental variables approaches and find a deterrent effect on crime due to passage of the law.¹ Marcet (2011) reaches a similar conclusion using a triple differences approach. Also using instrumental variables, Iyengar (2008) finds that the law reduced the probability of criminals committing additional crimes. By analyzing a specific population of criminals, Helland and Tabarrok (2007) find that the law reduced felony arrest rates. And, using a time series approach, Ramirez and Crano (2003) find a deterrent effect on nonviolent crimes and an incapacitation effect on both violent and nonviolent crimes due to the law's passage. Conversely, other papers find no evidence that the California law reduced crime. Using a panel data approach with fixed effects, Worrall (2004) finds no deterrent or incapacitation effects due to passage of the law. And, using a time series approach, Stolzenberg and D'alessio (1997) find that the law did not reduce crime rates.

However, fewer studies have examined how three strikes laws impacted crime outcomes in other states. One such multistate study is Chen (2008), who uses a panel data approach with fixed effects and finds that three strikes laws reduced crime.

¹Datta (2017) also claims to find an incapacitation effect.

On the other hand, Marvell and Moody (2001) find that three strikes laws increased homicides using a similar approach. While many of the California studies conduct analyses using county- or municipal-level data, very few 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 using a panel data approach with fixed effects. They find that three strikes laws increase homicide rates but have no effect on other crime rates.

Second, this paper improves on previous identification strategies used to analyze three strikes laws by restricting the analysis to within “border regions”.² This means that the sample is restricted to include only “border municipalities” (i.e., municipalities located in a county bordering another state with a different treatment status), and border-region-by-year fixed effects are used to ensure that the variation affecting estimates comes solely from within municipalities over time and from changes in treatment status within a border region and year. This approach takes 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 use similar techniques, such as Dube, Lester, and Reich (2010), who study the effects of minimum wages on employment using county-pair-by-year fixed effects and a sample of contiguous county pairs across state borders.

The rest of the paper is structured as follows. Section 2.2 provides background information on three strikes laws. Section 2.3 describes the data, and Section 2.4

²In this paper, a “border region” includes all municipalities located in counties that comprise a particular state border. For example, the Missouri/Kansas border region includes all municipalities (in either state) located in counties that comprise the Missouri/Kansas border.

discusses the empirical strategy. Section 2.5 outlines the main results, Section 2.6 discusses additional topics, and Section 2.7 concludes and suggests directions for future research.

2.2 Law Background

During the 1990s, 24 U.S. states passed some type of three strikes law (Clark, Henry, and Austin, 1997). These laws mandate increased sentences for criminals who commit two to four qualifying felonies (the exact number depends on the state). Figure 2.1 summarizes the status of three strikes laws for all U.S. states. States that are orange, blue, or yellow passed a three strikes law in the 1990s. States in gray 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 (e.g., murder, robbery, rape, aggravated assault) fall under the purview of state courts and are subject to state-level 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 constitute a strike (Clark, Henry, and Austin, 1997). Some states count other crimes as strikes as well, such as selling drugs to minors in California (Clark, Henry, and Austin, 1997). The consequences of repeat offending also vary among states. In 12 states, repeat offenders receive a mandatory life sentence without the possibility of parole (Clark, Henry, and Austin, 1997), while in others, they may receive a life sentence but with the possibility of pa-

role (Clark, Henry, and Austin, 1997). Regardless, in all 24 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, while Maryland requires four (Clark, Henry, and Austin, 1997).³

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).⁴ 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, the state 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,

³There 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 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.

⁴There may be one exception among states. 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 while Clark, Henry, and Austin (1997) claim it did not.

Henry, and Austin, 1997).⁵ Since such offenders would be most affected by the passage of three strikes laws, I restrict the outcome variables in my analyses to be measures of violent crime.

2.3 Data

The paper uses violent crime data from the Uniform Crime Reporting (UCR) Program. The UCR Program is compiled by the FBI and provides official data on reported criminal offenses in the United States. These data originate from municipalities with at least 10,000 residents and cover the time period from 1988 to 2000.

As defined by the UCR Program, violent crime consists of four offenses: murder and non-negligent manslaughter, rape, robbery, and aggravated assault. The UCR Program has precise definitions for what constitutes each offense, but reporting inconsistencies can occur. For example, numerous municipalities reported rape offenses incorrectly to the FBI during the 1985–2014 time period. Consequently, the UCR Program removed rape and violent crime counts from its database in the years where there was a reporting issue for the municipalities in question.⁶ Thus, 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 Program 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

⁵California, Florida, Indiana, Louisiana, South Carolina, and Washington are the six states that include some non-violent crimes as strikes (Clark, Henry, and Austin, 1997).

⁶This issue was especially pronounced in Illinois, where every single municipality had its rape and violent crime counts removed for the years 1985–2005.

to my adjusted version of violent crime that omits rape data.⁷

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 come from samples of U.S. census data compiled by IPUMS USA (Flood et al., 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 missing years. Additionally, for any county present in the UCR Program 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 datasets. Thus, the regressor values for counties appearing only in the UCR Program data do not change over time, which I account for by including a “missing regressor” dummy in the regressions that equals one if a particular county lacks regressor data. Doing this allows me to retain my full sample while ensuring that my estimates are only affected by regressor values that change over time.

Table 2.1 lists the data’s summary statistics. Statistics for the full sample are listed in the left-most column, and statistics for both the treated and control subsamples are in the next two columns. The total sample and each subsample are fairly similar across all metrics, although the control subsample has a slightly higher black share of the population (13.1%) compared to the treated subsample (9.6%). Note the high

⁷As 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, and these results are compiled in Appendix B.2.

standard deviations for the violent crime rate across all samples (502, 485, and 523 for total, treated, and control, respectively, compared to the means of 428, 444, and 406).⁸ These high standard deviations indicate a high degree of dispersion in the outcome data.

2.4 Methodology

2.4.1 Main Analysis

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

$$Y_{mcr,t} = \alpha + \beta T_{mt} + \gamma \mathbf{X}_{\mathbf{crt}} + \delta_m + \theta_{rt} + \epsilon_{mcr,t}, \quad (2.1)$$

where m indexes municipalities, r indexes border regions, t indexes years, and c indexes counties. $Y_{mcr,t}$ is the the number of violent crimes 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 a vector of county-level regressors that 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. δ_m are municipality fixed effects, θ_{rt} are border-region-by-year fixed effects, α is the intercept, and $\epsilon_{mcr,t}$ is the error term.

Some of the municipalities in the data have undergone changes to their crime

⁸The violent crime rate is defined as the number of violent crimes per 100,000 people in the UCR Program data.

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 equation (2.1) and the sample selection process are inspired by the analyses of minimum wage laws in both Dube, Lester, and Reich (2010) and Card and Krueger (1993). As previously mentioned, Card and Krueger (1993) examine a sample of fast-food restaurants on either side of the New Jersey/Pennsylvania border to analyze the effect of a New Jersey minimum wage law on employment. By restricting their sample to only include border-region restaurants, they obtain more accurate findings since any estimated effect is due to differences in law status (based on where the restaurants are located) and is not due to differing labor markets. I aim to accomplish something similar by restricting my sample to include only municipalities in counties that border a state with a different three strikes law treatment status.⁹

One 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, consider a municipality in Washington 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

⁹As a robustness check, I run my analysis using a sample containing all border municipalities. The results are in Appendix B.2 and are consistent with my main findings.

Idaho and Oregon both have different treatment statuses than Washington. On the other hand, consider an Arkansas municipality located in a county bordering both Missouri and Tennessee but 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 the 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 from variation within a border region and year due to differences in treatment status.

I cluster the confidence intervals from equation (2.1) at the treatment level (i.e., state level).¹⁰ In addition, I use the wild cluster bootstrap to construct the confidence intervals for the 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).¹¹ I choose to use the wild cluster bootstrap to construct the 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

¹⁰As a robustness check, I cluster the confidence intervals at the border-region level. The results appear in Appendix B.2 and are consistent with my main findings.

¹¹I use the recommended options in the `boottest` package in Stata to construct the wild cluster bootstrap confidence intervals (Roodman et al., 2019).

1993–1995, with the full study period ranging from 1988 to 2000), reducing concerns of such bias. In addition, Goodman-Bacon’s (2018) concerns do not apply to event studies. Thus, the fact that the event study results in Section 2.5 are similar to the results using the difference-in-differences specification provides additional support for minimal bias resulting from treatment timing heterogeneity.

2.4.2 Pre-Trend Analysis

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

$$Y_{mcr,t} = \alpha + \beta_{-5}T_{m,t}^{-5} + \beta_{-4}T_{m,t}^{-4} + \beta_{-3}T_{m,t}^{-3} + \beta_{-2}T_{m,t}^{-2} + \beta_0T_{m,t}^0 + \beta_1T_{m,t}^1 + \beta_2T_{m,t}^2 + \beta_3T_{m,t}^3 + \beta_4T_{m,t}^4 + \beta_5T_{m,t}^5 + \gamma\mathbf{X}_{crt} + \delta_m + \theta_{rt} + \epsilon_{mcr,t}, \quad (2.2)$$

where $T_{m,t}^0$ is an indicator variable set equal to one for the year t in which the state containing municipality m passed a three strikes law, $T_{m,t}^5$ is set equal to one five or more years after the passage of the law in the state containing municipality m , $T_{m,t}^4$ is set equal to one four years after the passage of the law, and so on. Similarly, $T_{m,t}^{-5}$ is set equal to one five or more years prior to the passage of a three strikes law in the state containing municipality m , $T_{m,t}^{-4}$ is set equal to one four years prior to the passage of the law, and so on. $T_{m,t}^{-1}$ is omitted from the model to serve as a baseline for comparison. $Y_{mcr,t}$, \mathbf{X}_{crt} , δ_m , θ_{rt} , α , and $\epsilon_{mcr,t}$ are as stated in Section 2.4.1. Evidence consistent with the parallel trends assumption holding would consist of the parameter estimates for $T_{m,t}^{-5}$, $T_{m,t}^{-4}$, $T_{m,t}^{-3}$, and $T_{m,t}^{-2}$ being close to zero and statistically insignificant.

Unfortunately, the pre-treatment parameter estimates are different from zero and statistically significant when I run equation (2.2) on the 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 to improve the pre-trends.¹² Specifically, I match on municipal-level violent crime rates one, two, three, four, and five years prior to treatment using 1:N nearest-neighbor propensity score matching (with replacement), enforcing common support, and with five control units matched to each treatment unit.¹³ After the matching process is completed, I run equations (2.1) and (2.2) on the matched sample.

There is disagreement in the matching literature regarding whether to match on pre-treatment outcomes to improve pre-trends for difference-in-differences analyses. Ryan et al. (2019) and Chabé-Ferret (2017) find that matching on pre-treatment outcomes works well in simulations. In particular, Ryan et al. (2019) show that the difference-in-differences estimator with matching has lower mean squared error when compared to the standard difference-in-differences estimator. Other papers find 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 status is very low in my analysis.¹⁴

¹²The matching is done via the MatchIt software in R (Ho et al., 2007; Stuart et al., 2011).

¹³Mueser, 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.

¹⁴In my sample, the Pearson correlation coefficient equals 0.054, indicating very low correlation between pre-treatment outcomes and treatment status.

2.5 Results

2.5.1 Border Municipality Analysis

Figure 2.2 shows the the pre-trend analysis for equation (2.2) run on the matched sample. The figure includes point estimates and 95% confidence intervals for the estimated treatment effect in each year within a 10-year window around the treatment year (specifically, 5 years before and after the treatment year). Estimates for the year immediately preceding treatment are omitted to serve as a baseline, and 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 equation (2.2) 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 2.2 are indeed close to zero and are statistically insignificant, providing evidence consistent with the parallel trends assumption holding.

Roth (2018) argues there is potential for bias from pre-testing and that many pre-tests are underpowered. While my analysis does suffer from being somewhat underpowered, I reduce the potential for bias in my pre-trend testing since my study period is fairly short (1988–2000), meaning 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, which also makes the pre-trends assumption more plausible.

The results from running equation (2.1) on my border municipality sample are

compiled in Table 2.2.¹⁵ The table includes point estimates, 95% wild cluster bootstrap confidence intervals clustered at the state level, and both two-tailed and lower-tailed p -values.

Columns (1) and (2) show results for equation (2.1) run on the non-matched sample and include border-region weights that give equal weight to each border region included in equation (2.1). Column (2) contains county-level regressors, while column (1) does not. Even though 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 the results should be interpreted with caution.

Columns (3) and (4) in Table 2.2 show results where municipalities are matched based on pre-treatment outcomes and then equation (2.1) 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 municipal-level 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 10% 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. That being said, my analysis is somewhat underpowered in that I cannot rule out the possibility of a negative effect. However, I can rule out a negative effect larger than

¹⁵Various robustness checks that slightly adjust my main specification are included in Appendix B.2. The results are robust to these specification changes.

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 2.2. Additionally, the fact that that my point estimates do not change significantly when adding in controls (i.e., moving from column (3) to (4)) provides support to my choice of specification as an identification strategy.

Note that the border municipality results in Table 2.2 include spillover effects. That is, if we believe criminals to be rational actors, we might expect the passage of a three strikes 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, the border municipality results are still positive. To disentangle such spillover effects from the law's real effect, I pursue an interior municipality analysis in Section 2.5.2.

2.5.2 Interior Municipality Analysis

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 equation (2.1) as my specification, but now the 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 2.5.1 do not change substantially when they are included.

To maintain the same specification across the border municipality and interior

municipality analyses, I assign a border region to each interior municipality. Doing this presents a problem since 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, assume 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 that, for the 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.

Figure 2.3 shows the pre-trend analysis for equation (2.2) run on the interior municipality matched sample, using the same matching procedure as in Section 2.5.1. This figure includes point estimates and 95% confidence intervals for the estimated treatment effect in each year within a 10-year window around the treatment year (specifically, 5 years before and after the treatment year). Estimates for the year immediately preceding treatment are omitted to serve as a baseline, and 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 equation (2.2) being close to zero and statistically insignificant. Similar to the the border municipality analysis, the interior

municipality analysis suffers from being underpowered, as can be seen by the somewhat wide confidence intervals. However, the pre-treatment coefficient estimates in Figure 2.3 are indeed close to zero and statistically insignificant, providing evidence consistent with the parallel trends assumption holding.

Table 2.3 shows the results from running equation (2.1) on the interior municipality sample. The table 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 equation (2.1) 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 included in equation (2.1). The point estimate for column (1) is statistically significant at the 5% level. However, as this unmatched version of my analysis does not satisfy the parallel trends assumption, the results should be interpreted with caution.

Column (2) in Table 2.3 displays results where I match municipalities based on pre-treatment outcomes and then run equation (2.1) on the matched sample. As in Section 2.5.1, 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 municipal-level 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 10% 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

underpowered 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 2.3.

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 2.5.1 to overestimate the negative effect of the laws on crime outcomes. Thus, the removal of such spillovers, via the interior municipality analysis, would result in less negative estimates. In fact, the results from Section 2.5.2 are statistically indistinguishable from Section 2.5.1, providing no evidence of spillover effects.

2.5.3 Falsification Tests

To test whether my results in Section 2.5.1 are spurious, I evaluate the effect of three strikes laws on outcomes that should be unaffected by the 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 2.5.1.

Table 2.4 displays the results from the falsification tests. The table 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 (results indicate that the 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%. Also, some statistically significant results are likely to occur by chance when undertaking multiple hypothesis testing. In summary, the falsification tests support my findings in Section 2.5.1, with the lone statistically significant result for black share of the population likely to have occurred by chance.

2.6 Additional Discussion

2.6.1 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 two to five years after the passage of said laws. California sentenced 26,074 people under its three strikes law as of December 31, 1996 (Clark, Henry, and Austin, 1997), by far 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 that they can no longer commit crimes) of three strikes laws may not be fully realized due to the 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. Thus, this paper's analysis of the effects of three strikes laws on violent crime rates provides effect estimates when three strikes implementation is heterogeneous.

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

2.6.2 Disparate Results in the Three Strikes Literature

In Section 2.5.1, 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 analyze California's three strikes law, whereas most that find that three strikes laws had no effect on crime outcomes

study three strikes laws in multiple states.¹⁶ In addition, many papers that analyze California's three strikes law find that crime reduction came via a deterrent effect; that is, harsher penalties deterred criminals (or potential criminals) from committing crimes.

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 et al., 2017; Wright, 2010) and thus do not consider the potential sentence when deciding 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 1990s, 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. Thus they do not view enhanced penalties as harshly as writers of three strikes laws might anticipate. Paternoster (2010) references interviews with arrested individuals who perceive a 5-year prison sentence to be only twice as severe as a 1-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 the 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

¹⁶Chen (2008) is an exception, as she undertakes a multistate analysis and finds that three strikes laws reduced crime.

offense for which they were imprisoned. These findings cast 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.

Although this paper provides evidence that three strikes laws have little to no impact on violent crime rates via a deterrent effect, it may be possible that three strikes laws could reduce crime via incapacitation effects as 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.¹⁷ Thus, incapacitation effects could exist but have yet to appear in the data. Furthermore, many three strikes papers focus exclusively on deterrent effects or do not separate deterrent effects from incapacitation effects, so the question of incapacitation effects in the context of three strikes laws remains relatively unstudied.¹⁸ On the other hand, three strikes sentencing was applied infrequently in most states, as mentioned in Section 2.6.1, which would consequently limit the scope of any incapacitation effects.

2.7 Conclusion and Future Work

This paper analyzes the effect of three strikes laws on violent crime rates using municipal-level data from the FBI. 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 no statisti-

¹⁷Disposition means the final outcome of a prosecution.

¹⁸Ramirez and Crano (2003), Worrall (2004), Datta (2017), Kovandzic, Sloan III, and Vieraitis (2004), and Chen (2008) are the papers that claim to estimate incapacitation effects of three strikes laws.

cal 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 2.6.2, many papers have studied whether three strikes laws deter 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), Datta (2017), Kovandzic, Sloan III, and Vieraitis (2004), and Chen (2008) would be welcome. Furthermore, although numerous papers have examined California's three strikes law, few have analyzed these laws in other states that frequently apply enhanced sentencing. Studying three strikes laws in these states would clarify whether frequently applying three strikes sentencing would reduce violent crime rates.

Nonetheless, my paper's results can encourage policymakers to act cautiously when deciding whether 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.

Table 2.1: Summary Statistics (Full Sample), Three Strikes Laws

	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 (number of violent crimes 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. The violent crime data come from Uniform Crime Reporting (UCR); all other data come from Flood et al. (2020).

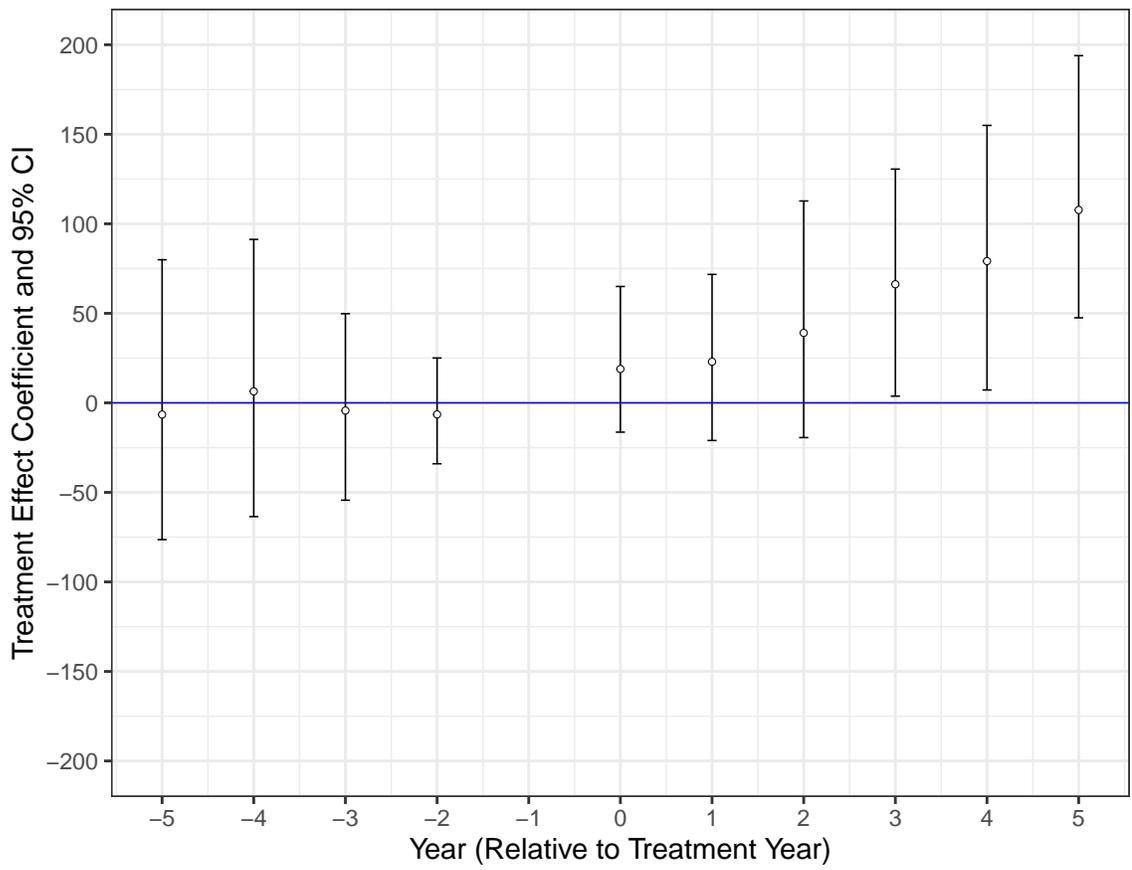


Figure 2.2: Border Municipality Pre-Trend Analysis, Three Strikes Laws

Table 2.2: 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. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

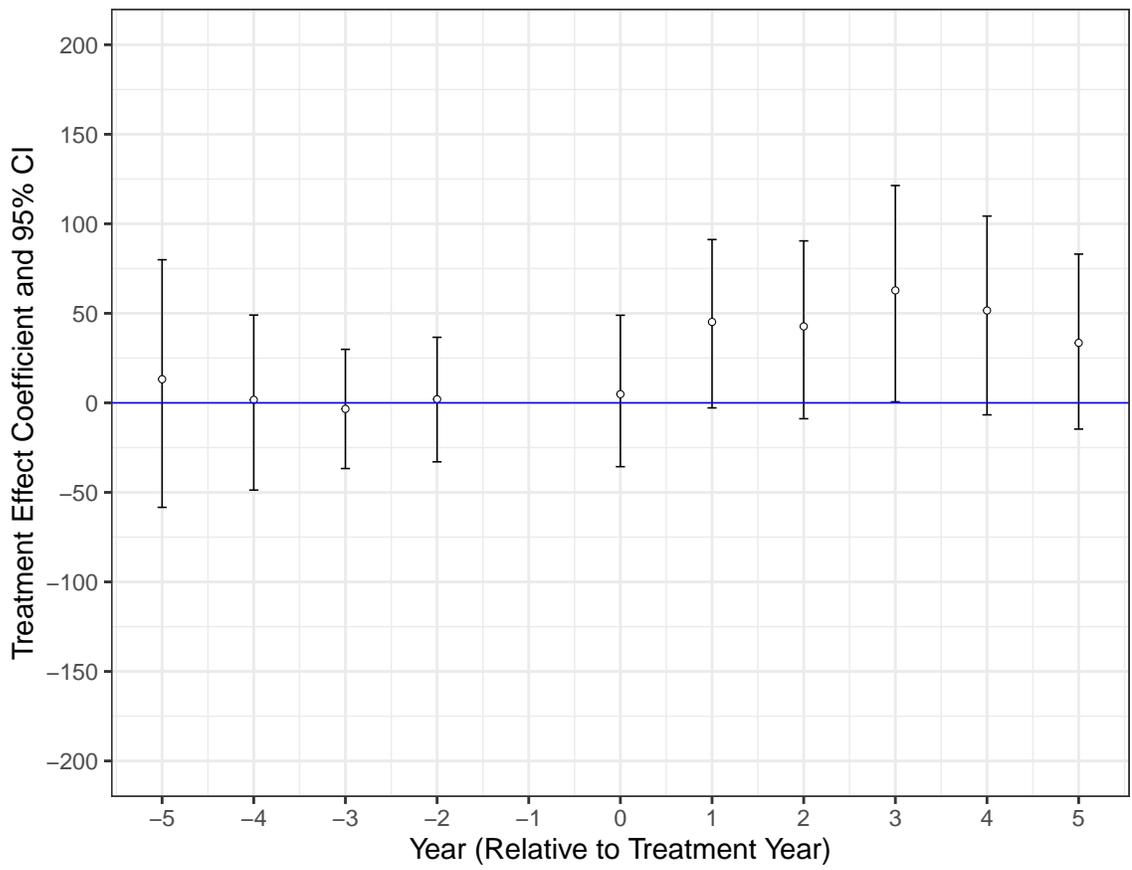


Figure 2.3: Interior Municipality Pre-Trend Analysis, Three Strikes Laws

Table 2.3: 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 Year 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. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Table 2.4: 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
<i>N</i>	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 passage of a three strikes law. Coefficients show the effect of policy on the variable indicated in the column and are measured 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 municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Chapter 3

The Impact of Legalizing Medical Marijuana on Violent Crime Rates

3.1 Introduction

Medical marijuana laws (MMLs) allow people with qualifying medical conditions to use, possess, and cultivate marijuana for medical purposes. One hypothesized benefit to MMLs is a reduction in violent crime rates due to decreased activity by drug trafficking organizations in jurisdictions where medical marijuana has been legalized. However, the impact of MMLs on violent crime rates remains under study.

In this paper, I study the effect of MMLs on violent crime rates using municipal-level crime data from the FBI. Comparing border municipalities with different treatment statuses using a difference-in-differences approach with a sample matched on pre-treatment outcomes, I find minimal evidence that MMLs increase violent crime rates. While I cannot rule out positive effects as large as 9.9% using a two-tailed test, I reject the hypothesis of an increase in violent crime rates due to passage of MMLs

at the 10% significance level using an upper-tailed test.

A number of studies in the medical marijuana literature find that MMLs decrease crime. For example, Huber III, Newman, and LaFave (2016) use a panel data approach with state and year fixed effects and find that MMLs reduced robberies, larcenies, and burglaries by 4-12%. Moreover, Shepard and Blackley (2016) use a panel data approach with state fixed effects to analyze MMLs in 11 states in the western United States and conclude that MMLs decrease violent crime rates. Finally, Gavrilova, Kamada, and Zoutman (2019) use a triple-differences approach to investigate the effect of MMLs in counties close to the Mexican border and find that MMLs reduce violent crime rates by 12.5%.

On the other hand, other studies find that MMLs do not have a noticeable effect on crime. For instance, using a difference-in-differences approach with city fixed effects and the synthetic control method, Chu and Townsend (2019) find no substantial effects (positive or negative) from MMLs on either violent or property crime. Morris et al. (2014) also investigate MMLs' effect on violent and property crime rates. They use a panel data approach with state fixed effects and find that MMLs did not increase either.

Recent papers also analyze the impact of recreational marijuana legalization on crime. Dragone et al. (2019) use a county-level difference-in-differences approach (along with spatial regression discontinuity) to analyze the effect of recreational marijuana legalization in Washington and Oregon, and find that legalization reduced rapes by 15% to 30% and thefts by 10% to 20%. Conversely, Wu, Wen, and Wilson (2021) find an increase in violent and property crime rates due to recreational marijuana legalization in Oregon using a difference-in-differences approach.

Other papers in the marijuana literature study topics related to marijuana use and

the impact of marijuana dispensaries. Investigating the former, and using a sample of arrestees in Arizona, Cheon, Decker, and Katz (2018) find that medical marijuana users are more likely to commit property and violent crimes. Investigating the latter, Hunt, Pacula, and Weinberger (2018) find no relationship between medical marijuana dispensaries and violent crime using a difference-in-differences approach in California counties.

While the effect of MMLs on crime has been well studied, this paper makes a contribution to the literature through its use of identification strategy. While other papers in the medical or recreational marijuana literature have leveraged the use of borders for their analysis (Dragone et al. (2019)) or analyzed the impacts of laws at the city- or county-level (Chu and Townsend (2019); Gavrilova, Kamada, and Zoutman (2019)), to my knowledge, no paper has investigated the impact of MMLs on violent crime rates by restricting the analysis to within “border regions” nationwide.¹

In my paper, restricting the analysis to within “border regions” implies that the sample is restricted to only include “border municipalities” (i.e., municipalities located in a county bordering another state), and border-region-by-year and municipality fixed effects are used to ensure that the variation affecting estimates comes from changes in treatment status within a border region in a given year and from within municipalities over time. This approach takes inspiration from Card and Krueger (1993) who use a sample of fast-food restaurants along the Pennsylvania/New Jersey border to evaluate the effect of a New Jersey minimum wage law on employment. Other papers use similar techniques: Dube, Lester, and Reich (2010) study the effects of minimum wages on employment using county-pair-by-year fixed effects and a

¹I define a “border region” to include all municipalities located in counties that constitute a particular state border. For example, the Nevada/Utah border region includes all municipalities (in either state) located in counties that comprise the Nevada/Utah border.

sample of contiguous county pairs across state borders, and Hofmann (2020) studies the effect of three strikes laws on violent crime rates using border-region-by-year fixed effects and a sample restricted to border municipalities.

The remainder of the paper is structured as follows. Section 3.2 provides background information on medical marijuana laws, and Section 3.3 describes the data. Section 3.4 discusses the empirical strategy, and Section 3.5 outlines the main results. Section 3.6 considers various mechanisms by which MMLs could impact violent crime rates, and Section 3.7 concludes and suggests directions for future research.

3.2 Background on Medical Marijuana Laws

From 1996–2003, medical marijuana laws (MMLs) were passed in eight states (the date of implementation is in parentheses): California (1996), Oregon (1998), Washington (1998), Alaska (1999), Maine (1999), Hawaii (2000), Colorado (2001), and Nevada (2001) (ProCon.org, 2021). These laws allow for the possession, use, and cultivation of marijuana for valid medical reasons. Figure 3.1 summarizes the status of MMLs for all U.S. states as of 2003. States that are non-gray passed MMLs at some point during the time period 1996–2003. States in gray did not pass MMLs during that time.

While MMLs allow for both the possession of marijuana for personal use and the possession of marijuana plants for cultivation, the amount of marijuana and the number of plants allowed varies from state to state (ProCon.org, 2021). However, no state allows the possession of more than 24 ounces or 24 plants at a given time. In addition, most of the legal supply of medical marijuana during the time period 1996–2003 likely came from home cultivation as opposed to medical marijuana dispensaries because,

as Gavrilova, Kamada, and Zoutman (2019) point out, most licensed dispensaries did not become active until the late 2000’s. The exception is California, where the first licensed dispensary opened in 1997 (Gavrilova, Kamada, and Zoutman, 2019).

The medical conditions that allow an individual to possess, use, and cultivate medical marijuana vary from state to state. However, there are many qualifying conditions common to all states, such as HIV/AIDS and cancer (ProCon.org, 2021). Furthermore, some of the qualifying conditions are sufficiently vague, such as chronic pain, as to allow for many people to qualify for medical use, possession, and cultivation under the laws (ProCon.org, 2021).

3.3 Data

This paper uses violent crime data from the Uniform Crime Reporting (UCR) program which is compiled by the FBI and provides data on reported criminal offenses in the United States. These data cover the time period from 1991–2003 and derive from municipalities with at least 10,000 residents. While states have passed medical marijuana laws (MMLs) since 2003, I choose the time period 1991–2003 for my study to reduce the concerns of bias from heterogeneous treatment timing articulated by Goodman-Bacon (2018).

The UCR Program defines violent crime as consisting of four offenses: murder and non-negligent manslaughter, rape, robbery, and aggravated assault, and has precise definitions for what constitutes each offense. However, reporting inconsistencies can occur. For example, many municipalities had substantial issues reporting rape offenses to the FBI during the time period 1985–2014. This largely had to do with local departmental definitions of rape diverging from FBI definitions. As a result, the rape

and violent crime counts were removed from the UCR program database in the years for which there was a reporting issue for the municipalities in question.

Since the reporting inconsistencies mostly affect the rape data, I follow the example of Hofmann (2020) and construct my own version of violent crime which does not include rape data in an effort to retain as many municipalities in my sample as possible. My “adjusted” measure of violent crime includes the three offenses in the UCR program definition of violent crime aside from rape: murder and non-negligent manslaughter, robbery, and aggravated assault. Hereinafter, violent crime refers to the adjusted measure that excludes rape offenses.

I incorporate regressors in my analyses that are deemed prognosticative of crime by the crime literature. County-level regressors are used since municipal-level regressors are not available for all municipalities in my sample. Similar to Hofmann (2020), the county-level regressors included in my regressions are: 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 come from samples of U.S. census data compiled by IPUMS USA (Flood et al., 2020).

I linearly interpolate to fill in the data for missing years since the regressor dataset from IPUMS does not include data for every year of my panel. Also, similar to Hofmann (2020), for any county present in the UCR Program 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 datasets. Thus, the regressor values for counties appearing only in the UCR Program data stay constant over time, and I account for this by including a “missing regressor” dummy in the regressions that equals one if a particular

county lacks regressor data. This approach allows me to keep my full sample while guaranteeing that my estimates are only affected by regressor values that change over time.

Table 3.1 contains summary statistics for the matched sample. Statistics for the full matched sample are given in the left-most column, while statistics for the treated and control sub-samples are given in the next two columns. The treated and control sub-samples are fairly similar across most metrics. Some exceptions include black and Hispanic shares of the population, where the treated sub-sample has a higher share of the latter (24% of the population versus 10%) and the the control sub-sample has a higher share of the former (11.4% of the population versus 6%). However, this should not cause concern as the statistics in Table 3.1 are for the full matched sample while my analyses occur within border regions.²

3.4 Methodology

3.4.1 Main Analysis

My main specification is the following difference-in-differences regression, which is similar to the analysis in Hofmann (2020):

$$Y_{mcr t} = \lambda + \alpha T_{mt} + \psi \mathbf{X}_{\mathbf{crt}} + \tau_m + \omega_{rt} + \epsilon_{mcr t}, \quad (3.1)$$

where m indexes municipalities, r indexes border regions, t indexes years, and c indexes counties. $Y_{mcr t}$ is the the number of violent crimes per 100,000 people, and T_{mt} is the treatment indicator variable. T_{mt} equals one in all years including and following

²The assumption being that characteristics between treated and control municipalities will be more similar within a given border region.

the year that the state containing municipality m legalized medical marijuana. T_{mt} equals zero in all years prior to the state containing municipality m legalizing medical marijuana. \mathbf{X}_{crt} is a vector of county-level regressors that 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. τ_m are municipality fixed effects, ω_{rt} are border-region-by-year fixed effects, λ is the intercept, and $\epsilon_{mcr t}$ is the error term.

Akin to Hofmann (2020), I specify the municipality fixed effects as municipality-by-reporting-regime fixed effects since some municipalities have undergone changes to their crime reporting structures over time and this can affect measured outcomes. For example, if Portland, Oregon changed its reporting structure in 1996, pre-1996 Portland and post-1996 Portland would be considered separate municipalities in my analysis.

Card and Krueger (1993), Dube, Lester, and Reich (2010), and Hofmann (2020) inspire this paper's sample selection process and the border-region-by-year fixed effects in equation (3.1). By restricting their sample to include only restaurants on either side of the New Jersey/Pennsylvania border, Card and Krueger (1993) obtain more accurate findings since any estimated effect of the New Jersey minimum wage law on employment is due to differences in law status (based on where the restaurants are located) and is not due to differing labor markets. By restricting my sample to include only municipalities located in counties that border another state, I hope to accomplish something similar in the context of medical marijuana laws (MMLs).

Some municipalities are located in counties that border multiple states. Since I cannot assign two or more border regions to a given municipality, similar to the

approach in Hofmann (2020), these municipalities are assigned the border region that includes the nearest border state with a differing treatment status. Furthermore, once the sample is restricted as stated previously, the border-region fixed effects then restrict the analysis further so that only municipalities *within* a border region are being compared (Dube, Lester, and Reich, 2010). Consequently, the estimated effects of MMLs will come from variation within a border region and year due to differences in treatment status and within municipalities over time.

The confidence intervals for the parameters of interest from equation (3.1) are constructed using the wild cluster bootstrap and clustered at the state level.³ The wild cluster bootstrap is used because typical cluster-robust standard errors can be inaccurate in analyses with a small number of clusters (Cameron, Gelbach, and Miller, 2008). My analysis exceeds the typical cutoff thought necessary to use conventional cluster-robust standard errors (Angrist and Pischke (2008) claim the cutoff to be 42 clusters), but I choose to be cautious and use the wild cluster bootstrap nonetheless.⁴

Bias can arise in difference-in-differences analyses due to treatment timing heterogeneity (Goodman-Bacon, 2018). Such concerns are reduced in this paper since the panel is constructed such that treatment timing does not occur near the beginning or end of the study period for any treated states. In addition, the similarity of the event study and difference-in-differences results in Section 3.5 seems to imply limited impact from bias due to treatment timing heterogeneity since such bias does not apply to event studies (Goodman-Bacon, 2018).

³The recommended options in the `boottest` package in Stata are used to construct the wild cluster bootstrap confidence intervals (Roodman et al., 2019).

⁴My main analysis has 43 clusters.

3.4.2 Pre-Trend Analysis

To test the parallel trends assumption, one of the key underlying assumptions of a differences-in-differences model, the following event study specification is used:

$$Y_{mcr,t} = \lambda + \alpha_{-5}T_{mt}^{-5} + \alpha_{-4}T_{mt}^{-4} + \alpha_{-3}T_{mt}^{-3} + \alpha_{-2}T_{mt}^{-2} + \alpha_0T_{mt}^0 + \alpha_1T_{mt}^1 + \alpha_2T_{mt}^2 + \psi\mathbf{X}_{\mathbf{crt}} + \tau_m + \omega_{rt} + \epsilon_{mcr,t}, \quad (3.2)$$

where T_{mt}^0 is an indicator variable set equal to one for the year t in which the state containing municipality m legalized medical marijuana, T_{mt}^2 is set equal to one two or more years after the legalization of medical marijuana in the state containing municipality m , and T_{mt}^1 is set equal to one in the year immediately following when the state containing municipality m legalized medical marijuana. T_{mt}^{-5} is set equal to one five or more years prior to the legalization of medical marijuana in the state containing municipality m , T_{mt}^{-4} is set equal to one four years prior to legalization, and so on. T_{mt}^{-1} is omitted from the model to serve as a baseline for comparison. $Y_{mcr,t}$, $\mathbf{X}_{\mathbf{crt}}$, τ_m , ω_{rt} , λ , and $\epsilon_{mcr,t}$ are as stated in Section 3.4.1. 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.

When I run equation (3.2) on the full sample, the pre-treatment parameter estimates are different from zero and statistically significant which is evidence consistent with poor parallel pre-trends. Thus, I attempt to improve the pre-trends by employing propensity score matching on pre-treatment outcomes.⁵ Similar to Hofmann (2020), I match on municipal-level violent crime rates one, two, three, four, and five years prior to treatment. I use 1:N nearest-neighbor propensity score matching (with

⁵The MatchIt software in R is used to do the matching (Ho et al., 2007; Stuart et al., 2011).

replacement), enforcing common support, with ten control units matched to each treatment unit.⁶ After the matching process is completed, equations (3.1) and (3.2) are run on the matched sample.

The matching literature disagrees on whether to match on pre-treatment outcomes to improve pre-trends for difference-in-differences analyses. Daw and Hatfield (2018) find that matching on pre-treatment outcomes can lead to bias when the pre-treatment outcomes are correlated with treatment status. However, the correlation between pre-treatment outcomes and treatment status is low in my analysis (the Pearson correlation coefficient equals only 0.099) which minimizes concerns of such bias. Conversely, Ryan et al. (2019) and Chabé-Ferret (2017) find that matching on pre-treatment outcomes works well in simulations. Specifically, Ryan et al. (2019) show that the difference-in-differences estimator with matching has lower mean squared error when compared to the standard difference-in-differences estimator.

3.5 Results

3.5.1 Border Municipality Analysis

The pre-trend analysis for equation (3.2) run on the matched sample can be seen in Figure 3.2. The figure includes point estimates and 95% confidence intervals for the estimated treatment effect in each year, from five years prior to treatment to two years after treatment. I omit estimates for the year immediately preceding treatment so they can serve as a baseline. Y-axis units are a rate per 100,000 people.

⁶Matching only one control to each treatment (with replacement) can increase standard errors due to sampling error, whereas using five or ten nearest neighbors does not increase standard errors to the same extent (Mueser, Troske, and Gorislawsky, 2007). So, in an attempt to avoid the sampling error problem, I choose to use ten nearest neighbors.

Pre-treatment coefficient estimates in equation (3.2) that are close to zero and statistically insignificant would be evidence consistent with the parallel trends assumption holding. The confidence intervals are wide, implying an underpowered analysis, but the pre-treatment coefficient estimates in Figure 3.2 are still statistically insignificant and somewhat close to zero. Thus, the results are consistent with the parallel trends assumption holding.

Some concerns with pre-testing are articulated by Roth (2018) who claims that many pre-tests are underpowered and there is potential for bias from pre-testing. Despite my analysis being underpowered, I attempt to reduce the bias from pre-testing by employing a short sample period of only 13 years (1991–2003). This means there is less time for other events to occur that could bias the event study results prior to treatment. Furthermore, matching on pre-treatment outcomes manually ensures that my treatment and control samples are similar in both trends and levels prior to treatment, which makes the pre-trends assumption more plausible.

The tables in this paper are organized similarly to those in Hofmann (2020). Table 3.2 contains results from running equation (3.1) on the border municipality sample. The table includes point estimates, 95% wild cluster bootstrap confidence intervals clustered at the state level, and both two-tailed and upper-tailed p -values.

Columns (1) and (2) display results from running equation (3.1) on the non-matched sample. This analysis includes border-region weights that give equal weight to each border region included in equation (3.1). Column (2) contains county-level regressors while column (1) does not. Neither point estimate in either column is statistically significant at all common significance levels, although I can reject the hypothesis of a positive effect at the 10% significance level for column (1). However, this version of my analysis does not satisfy the parallel trends assumption so the

results should be interpreted with caution.

Columns (3) and (4) in Table 3.2 display results for the analysis where equation (3.1) is run on the matched sample after municipalities are matched based on pre-treatment outcomes. The matching is done via 1:N nearest-neighbor propensity score matching (with replacement), enforcing common support, with ten controls matched to each treatment. As in Hofmann (2020), the only variables used for the matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment. The only difference between the specifications in columns (3) and (4) is that column (4) includes county-level regressors while column (3) does not.

Neither point estimate in either column is statistically significant at common significance levels. My preferred specification in column (4), with municipality fixed effects, border-region-by-year fixed effects, county-level controls, and matching weights, indicates that legalizing medical marijuana reduces the violent crime rate by 55.5 violent crimes per 100,000 people. This translates to a 10.9% reduction in the violent crime rate. However, I cannot rule out a positive effect as large as 9.9% due to my analysis being underpowered with somewhat wide confidence intervals. However, I can reject the hypothesis of a positive effect at the 10% significance level, shown via the p -value for my upper-tailed test in Table 3.2. Furthermore, my choice of specification as an identification strategy is bolstered by the fact that my point estimates do not change significantly when adding in controls (i.e., moving from column (3) to (4)).

Gavrilova, Kamada, and Zoutman (2019) point out that there exists a positive correlation between increased activity of drug-trafficking organizations in a jurisdiction and violence in said jurisdiction. They also mention that legalizing medical marijuana can potentially reduce profits for drug-trafficking organizations, thereby

causing them to reduce their presence and activity in states with medical marijuana laws (MMLs). If such reductions in profits were to occur in states with MMLs, we might expect drug-trafficking organizations to move their activities to states without MMLs. If this movement of criminal activity from treated to non-treated states exists, the resulting “spillover effects” in criminal activity would be captured in my border municipality analysis. Furthermore, the direction of bias from any spillover effects is likely to be negative. I undertake an interior municipality analysis in Section 3.5.2 to disentangle such spillover effects from the law’s true effect.

3.5.2 Interior Municipality Analysis

I analyze the effect of medical marijuana laws (MMLs) on violent crime rates using a sample of interior municipalities to disentangle the possible effect of drug-trafficking organizations moving criminal activity across borders from treated to non-treated states (i.e., spillover effects). I use equation (3.1) as my specification, but now the sample consists of interior municipalities as opposed to border municipalities. Also, county-level controls are omitted for the interior county analysis since my results in Section 3.5.1 do not change substantially when they are included.

Similar to the approach in Hofmann (2020), I assign a border region to each interior municipality to maintain the same specification across the border municipality and interior municipality analyses. Since interior municipalities, by definition, are located in counties that do not border other states, I randomly assign border regions to interior municipalities based on the frequency that border municipalities appear in a given border region. For example, Oregon borders four states: Washington, Nevada, Idaho, and California. Thus, each Oregon border county has one of three border regions:

Oregon–Washington (ORWA), Oregon–Nevada (ORNV), Oregon–Idaho (ORID), and Oregon–California (ORCA).

In this example, let’s assume that the border-region distribution for Oregon border municipalities is as follows: 10% of Oregon border municipalities are located in the ORCA border region, 30% in the ORNV border region, 30% in the ORWA border region, and 30% in the ORID border region. This means that 10% of interior Oregon municipalities would be randomly assigned the ORCA border region, 30% would be randomly assigned the ORNV border region, 30% would be randomly assigned the ORWA border region, and 30% would be randomly assigned the ORID border region. Once the border regions have been distributed, (3.1) and (3.2) are run on either the matched or non-matched versions of the interior municipality sample.

The pre-trend analysis for equation (3.2) run on the matched sample can be seen in Figure 3.3. The figure includes point estimates and 95% confidence intervals for the estimated treatment effect in each year, from five years prior to treatment to two years after treatment. I omit estimates for the year immediately preceding treatment so they can serve as a baseline. Y-axis units are a rate per 100,000 people.

Pre-treatment coefficient estimates in equation (3.2) that are close to zero and statistically insignificant would be evidence consistent with the parallel trends assumption holding. Similar to the the border municipality analysis, the confidence intervals are wide, implying an underpowered analysis; but the pre-treatment coefficient estimates in Figure 3.3 are still statistically insignificant and somewhat close to zero. Thus, the results are consistent with the parallel trends assumption holding.

As in Section 3.5.1, the table for the interior municipality analysis is organized similarly to the tables in Hofmann (2020). Table 3.3 contains results from running equation (3.1) on the interior municipality sample. The table includes point estimates,

95% wild cluster bootstrap confidence intervals clustered at the state level, and both two-tailed and upper-tailed p -values. Column (1) displays results for equation (3.1) 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 included in equation (3.1). The point estimate for column (1) is statistically significant at the 10% level. However, this version of my analysis does not satisfy the parallel trends assumption so the results should be interpreted with caution.

Column (2) in Table 3.3 displays results for the analysis where equation (3.1) is run on the matched sample after municipalities are matched based on pre-treatment outcomes. As in Section 3.5.1, the matching is done via 1:N nearest-neighbor propensity score matching (with replacement), enforcing common support, with ten controls matched to each treatment. Like in Hofmann (2020), the only variables used for the matching are municipal-level 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 5% level and indicates that legalizing medical marijuana reduces the violent crime rate by 30.9 violent crimes per 100,000 people. This translates to a 6.9% reduction in the violent crime rate. In addition, I can rule out a positive effect at the 5% significance level, and I can also reject the hypothesis of a positive effect at the 5% significance level, shown via the p -value for my upper-tailed test in Table 3.3.

I would expect my estimates in Section 3.5.1 to overestimate the negative effect of MMLs on crime outcomes if the passage of MMLs in treatment states caused crime to spill over into non-treatment states due to relocation of drug trafficking organizations. As such, removing the spillovers via the interior municipality analysis would result

in less negative estimates. While the border municipality analysis allows for larger negative effects than the interior municipality analysis, it also allows for positive effects while the interior municipality does not. Thus, given the similarity of the point estimates in Table 3.2 and Table 3.3 and the fact that the border municipality analysis allows for both more negative and less negative estimates than the interior municipality analysis, I conclude that the results in each analysis are fairly similar and do not provide compelling evidence of spillover effects.

3.6 Discussion of Possible Mechanisms

Medical marijuana laws (MMLs) have the potential to impact violent crime rates in a number of ways. For example, there is a connection between drug use and criminal activity (Bennett, Holloway, and Farrington, 2008). Specifically in regards to marijuana, Cheon, Decker, and Katz (2018) conclude that medical marijuana users are more likely to commit violent crimes. If their conclusions are correct and, in addition, MMLs increase the number of medical marijuana users, MMLs could cause an increase in violent crime rates.

Another way in which MMLs could impact violent crime rates is through a reorganization of policing priorities. If, due to passage of MMLs, police allocate resources away from drug enforcement and towards other policing priorities, we could see a decrease in other crimes. For example, Adda, McConnell, and Rasul (2014) find that depenalizing possession of small quantities of marijuana in London resulted in a decrease in crime as police reallocated resources away from drug enforcement.

While there are a number of mechanisms by which MMLs could influence violent crime rates, the impact of MMLs on drug trafficking organizations is especially

compelling. Gavrilova, Kamada, and Zoutman (2019) investigate this mechanism specifically and find that MMLs in counties close to the Mexican border reduce violent crime rates by 12.5%. This reduction in violent crime rates could be caused by drug trafficking organizations decreasing their presence in Mexican border counties due to lower profits induced by MMLs (Gavrilova, Kamada, and Zoutman, 2019).

My results in Section 3.5 do not rule out small-to-medium positive effects, but the statistical evidence allows for the possibility of large negative effects as well. If MMLs lead to such decreases in violent crime, two potential mechanisms driving the results could be two mentioned in previous paragraphs: a reallocation of policing priorities or a reduction in criminal activity from drug trafficking organizations. Interestingly, Gavrilova, Kamada, and Zoutman (2019) rule out the possibility of the former in their paper as they find that arrests for illicit drug offenses actually increased after passage of MMLs. Consequently, while I cannot conclude anything definitively, drug-trafficking organizations reducing their activities due to MMLs decreasing their profits could be a possibility for a potential decrease in violent crime rates.

3.7 Conclusion and Future Work

In this paper, I analyze the effect of legalizing medical marijuana on violent crime rates using municipal-level data from the FBI. I compare violent crime rates in border municipalities using a difference-in-differences specification with a sample matched on pre-treatment outcomes. My results provide little statistical justification that medical marijuana laws (MMLs) increase violent crime rates. I cannot rule out positive effects up to 9.9% using a two-tailed test, but I can reject the hypothesis that MMLs increase violent crime rates at the 10% significance level using an upper-tailed test.

The effects of MMLs on crime have been well-studied, but new states continue to pass MMLs frequently. For instance, Mississippi and South Dakota recently passed MMLs in 2020, and Virginia passed a MML in 2021 (ProCon.org, 2021). New studies that incorporate the effects of these newer laws on crime would be welcome.

In addition, additional research analyzing laws that legalize recreational marijuana would be informative. There has been some preliminary work in this area, such as Dragone et al. (2019) and Wu, Wen, and Wilson (2021) but, since recreational marijuana legalization is so recent, future work that looks at the long-term effects of recreational marijuana legalization on crime would be interesting.

In summary, the results of this paper reinforce much of the scholarly work on MMLs and violent crime. That is, there is minimal statistical evidence that MMLs lead to large increases in violent crime, and policymakers should feel relatively confident in the knowledge that passing a MML will likely not induce a large increase in violent crime rates. In addition, there is some support for the view that MMLs produce at least a modest decline in violent crime.

Table 3.1: Summary Statistics (After Matching), Medical Marijuana

	Total Mean (SD)	Treated Mean (SD)	Control Mean (SD)
Violent Crime Rate	507 (507)	567 (428)	493 (523)
Male	48.8 (0.7)	49.7 (0.6)	48.6 (0.6)
Black	10.4 (6.7)	6.0 (3.7)	11.4 (6.8)
Hispanic	12.6 (11.8)	24.0 (15.9)	10.0 (8.7)
18-44 Years Old	40.1 (2.5)	40.8 (2.2)	39.9 (2.6)
No HS degree	37.6 (4.5)	40.5 (5.7)	36.9 (3.9)
In poverty	10.6 (3.7)	12.4 (3.2)	10.2 (3.7)
Unemployed	6.3 (1.5)	7.1 (1.4)	6.1 (1.5)
N	6,305	1,170	5,135

Violent Crime is a rate (number of violent crimes 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 come from Uniform Crime Reporting (UCR); all other data come from Flood et al. (2020).

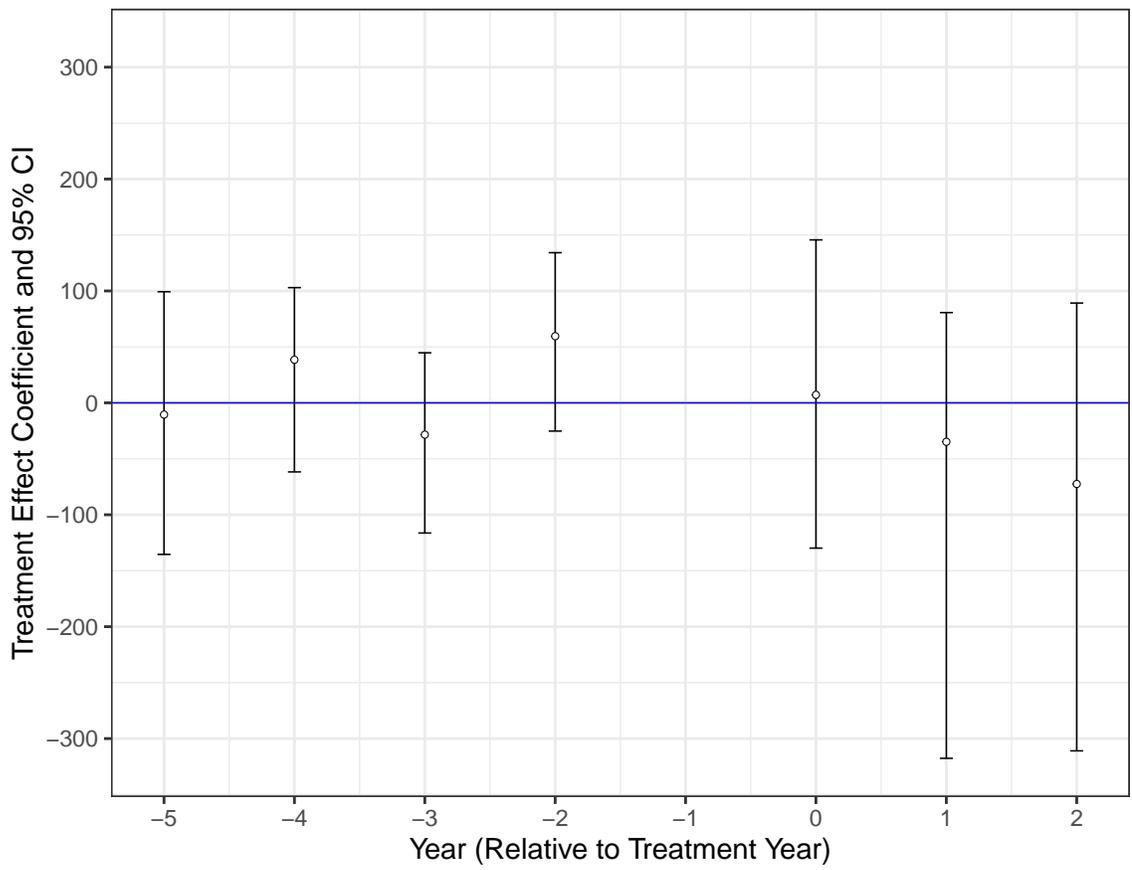


Figure 3.2: Border Municipality Pre-Trend Analysis, Medical Marijuana

Table 3.2: Effects of Medical Marijuana Legalization on Adjusted Violent Crime Rates, Border Municipality Analysis

	(1)	(2)	(3)	(4)
policy	-61.3 [-152.3, 4.7]	-54.4 [-168.1, 20.4]	-65.5 [-271.6, 51.3]	-55.5 [-223.5, 49.9]
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
<i>N</i>	16,848	16,848	6,305	6,305
Mean Dependent Variable	403	403	507	507
No. of clusters	48	48	43	43
P-value ($H_0 : \beta = 0$)	0.100	0.208	0.382	0.178
P-value ($H_0 : \beta > 0$)	0.050*	0.115	0.168	0.086*

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a law that legalized medical marijuana. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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 ten controls matched to each treatment. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

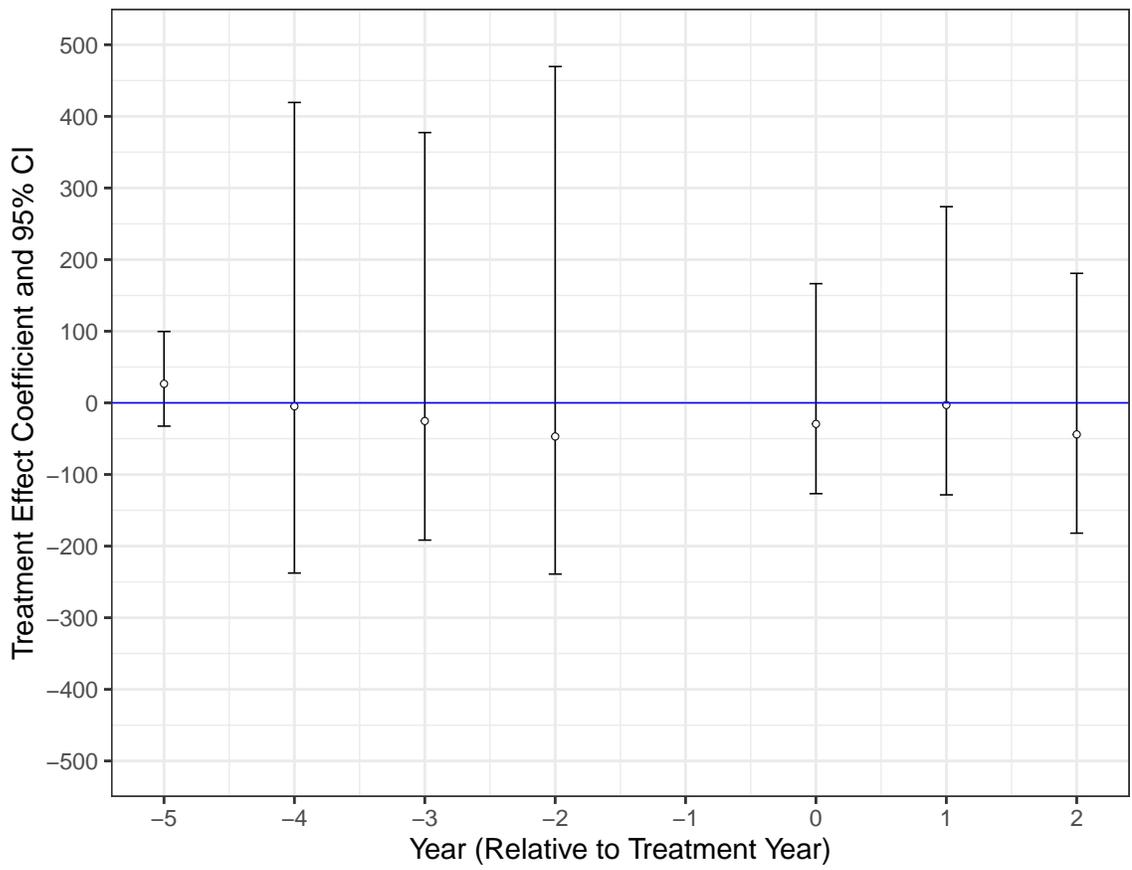


Figure 3.3: Interior Municipality Pre-Trend Analysis, Medical Marijuana

Table 3.3: Effects of Medical Marijuana Legalization on Adjusted Violent Crime Rates, Interior Municipality Analysis

	(1)	(2)
policy	-65.9 [-118.8, 27.2]	-30.9 [-54.9, -9.9]
Municipality FE	X	X
Border-Region x Year FE	X	X
County Controls		
Border-Region Weights	X	
Matching Weights		X
N	26,013	15,834
Mean Dependent Variable	452	449
No. of clusters	43	41
P-value ($H_0 : \beta = 0$)	0.072*	0.018**
P-value ($H_0 : \beta > 0$)	0.025**	0.011**

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a law that legalized medical marijuana. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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 ten controls matched to each treatment. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Appendix A

Appendix for Chapter 1

A.1 Proof of Theorem 5

Proof. Define $\frac{k+\epsilon}{n+1} \equiv u^l(\alpha)$. By Lemma 3 from Goldman and Kaplan (2017),

$$u^l(\alpha) = \tau - n^{-1/2} z_{1-\alpha} \sqrt{\tau(1-\tau)} - \frac{(2\tau-1)}{6n} (z_{1-\alpha}^2 + 2) + O(n^{-3/2}).$$

As $u^l(\alpha)$ is defined as the u that results in CP of $1-\alpha$, by extension, $u^l(\alpha - cn^{-1/2})$ is the u that results in CP of $1-\alpha + cn^{-1/2}$. I now compute the difference between $u^l(\alpha)$ and $u^l(\alpha - cn^{-1/2})$.

$$\begin{aligned} & u^l(\alpha) - u^l(\alpha - cn^{-1/2}) \\ & \quad \overbrace{= \tau - n^{-1/2} z_{1-\alpha} \sqrt{\tau(1-\tau)} - \frac{(2\tau-1)}{6n} (z_{1-\alpha}^2 + 2) + O(n^{-3/2})}^{\text{by Lemma 3 from Goldman and Kaplan (2017)}} \\ & \quad \overbrace{- [\tau - n^{-1/2} z_{1-\alpha+cn^{-1/2}} \sqrt{\tau(1-\tau)} - \frac{(2\tau-1)}{6n} (z_{1-\alpha+cn^{-1/2}}^2 + 2) + O(n^{-3/2})]}^{\text{by Lemma 3 from Goldman and Kaplan (2017)}}. \end{aligned}$$

By Taylor expansion of $z_{1-\alpha+cn^{-1/2}}$,

$$\begin{aligned}
& \tau - n^{-1/2} z_{1-\alpha} \sqrt{\tau(1-\tau)} - \frac{(2\tau-1)}{6n} (z_{1-\alpha}^2 + 2) + O(n^{-3/2}) \\
& \quad - [\tau - n^{-1/2} Z_{1-\alpha+cn^{-1/2}} \sqrt{\tau(1-\tau)} - \frac{(2\tau-1)}{6n} (z_{1-\alpha+cn^{-1/2}}^2 + 2) + O(n^{-3/2})] \\
& = -\frac{1}{\sqrt{n}} z_{1-\alpha} \sqrt{\tau(1-\tau)} - \frac{(2\tau-1)}{6n} (z_{1-\alpha}^2 + 2) + O(n^{-3/2}) \\
& \quad + \frac{1}{\sqrt{n}} [z_{1-\alpha} + \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} + O(n^{-1})] \sqrt{\tau(1-\tau)} \\
& \quad + \frac{(2\tau-1)}{6n} [(z_{1-\alpha} + \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} + O(n^{-1}))^2 + 2] - O(n^{-3/2}) \\
& = -\frac{(2\tau-1)}{6n} (z_{1-\alpha}^2 + 2) + O(n^{-3/2}) + n^{-1/2} \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} \sqrt{\tau(1-\tau)} \\
& \quad + n^{-1/2} O(n^{-1}) \sqrt{\tau(1-\tau)} + \frac{(2\tau-1)}{6n} [(z_{1-\alpha} + \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} + O(n^{-1}))^2 + 2] - O(n^{-3/2}).
\end{aligned}$$

By expansion of $(z_{1-\alpha} + \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} + O(n^{-1}))^2$,

$$\begin{aligned}
& -\frac{(2\tau-1)}{6n} (z_{1-\alpha}^2 + 2) + O(n^{-3/2}) + n^{-1/2} \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} \sqrt{\tau(1-\tau)} \\
& \quad + n^{-1/2} O(n^{-1}) \sqrt{\tau(1-\tau)} + \frac{(2\tau-1)}{6n} [(z_{1-\alpha} + \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} + O(n^{-1}))^2 + 2] - O(n^{-3/2}) \\
& = -\frac{(2\tau-1)}{6n} (z_{1-\alpha}^2 + 2) + O(n^{-3/2}) + n^{-1/2} \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} \sqrt{\tau(1-\tau)} \\
& \quad + n^{-1/2} O(n^{-1}) \sqrt{\tau(1-\tau)} + \frac{(2\tau-1)}{6n} [(z_{1-\alpha}^2 + 2z_{1-\alpha} O(n^{-1}) + \\
& \quad + (\frac{cn^{-1/2}}{\phi(z_{1-\alpha})})^2 + O(n^{-1})^2 + 2 + \frac{2z_{1-\alpha} cn^{-1/2}}{\phi(z_{1-\alpha})} + \frac{2cn^{-1/2} O(n^{-1})}{\phi(z_{1-\alpha})}] - O(n^{-3/2}) \\
& = O(n^{-3/2}) + \frac{c\sqrt{\tau(1-\tau)}}{n\phi(z_{1-\alpha})} + n^{-1/2} O(n^{-1}) \sqrt{\tau(1-\tau)} + 2\frac{(2\tau-1)}{6n} z_{1-\alpha} O(n^{-1}) \\
& \quad + \frac{c^2(2p-1)}{6n^2\phi(z_{1-\alpha})^2} + \frac{(2p-1)O(n^{-1})^2}{6n} + 2\frac{2p-1}{6n} \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} z_{1-\alpha} \\
& \quad + 2\frac{2p-1}{6n} \frac{cn^{-1/2}}{\phi(z_{1-\alpha})} O(n^{-1}) - O(n^{-3/2})
\end{aligned}$$

$$\begin{aligned}
&= O(n^{-3/2}) + \frac{c\sqrt{\tau(1-\tau)}}{n\phi(z_{1-\alpha})} + O(n^{-3/2}) + O(n^{-2}) + O(n^{-2}) + O(n^{-3}) + O(n^{-3/2}) \\
&\quad + O(n^{-5/2}) - O(n^{-3/2}) \\
&\doteq \frac{c\sqrt{\tau(1-\tau)}}{n\phi(z_{1-\alpha})}
\end{aligned}$$

□

Appendix B

Appendix for Chapter 2

B.1 Matching Summary Statistics

Additional summary statistics for my sample after matching are included in Table B.1.

B.2 Robustness Checks

Table B.2 includes results from running equation (2.1) on a matched sample that originated from a sample including all border municipalities. In other words, the initial sample was not restricted to include only border municipalities with different treatment statuses or timings, in contrast to my results in Section 2.5.1. As can be seen in Table B.2, the results are robust to this specification change.

Table B.3 includes results from running equation (2.1) on a matched sample with the wild cluster bootstrap confidence intervals clustered at the border-region level. As can be seen in Table B.3, the results are robust to this specification change.

Table B.4 includes results from running equation (2.1) 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 2.2 except for the additional use of a caliper set equal to 0.05.¹ As can be seen in Table B.4, the results are robust to this specification change.

Table B.5 includes results from running equation (2.1) on a matched sample where only municipalities from border regions comprised of one treatment state and one control state are included. In other words, I no longer include municipalities from border regions comprised of two treatment states with differing treatment timings. As can be seen in Table B.5, the results are robust to this specification change.

Table B.6 includes results from running equation (2.1) 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 B.6, the results are robust to this specification change.

Table B.7 includes results from running equation (2.1) on a matched sample where the sample is restricted to the six states (California, Florida, Georgia, Nevada, South Carolina, and Washington) that sentenced more than 100 criminals using provisions in their respective three strikes legislation in the two to five years after passage of said laws (Prison Policy Initiative, 1998). Unfortunately, the analysis is significantly underpowered due to small sample sizes, so the results are inconclusive.

¹Per 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 B.1: Summary Statistics (After Matching), Three Strikes Laws

	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 (number of violent crimes 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. The violent crime data come from Uniform Crime Reporting (UCR); all other data come from Flood et al. (2020).

Table B.2: Effects of Three Strikes Laws on Violent Crime Rates (Sample Includes All Border Municipalities)

	(1)	(2)	(3)	(4)
policy	20.3 [-26.6, 70.6]	27.7 [-21.6, 81.1]	57.9 [-37.3, 112.4]	56.4 [-15.5, 121.1]
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
<i>N</i>	16,848	16,848	10,764	10,764
Mean Dependent Variable	411	411	410	410
No. of clusters	48	48	44	44
P-value ($H_0 : \beta = 0$)	0.437	0.308	0.170	0.106
P-value ($H_0 : \beta < 0$)	0.231	0.169	0.080*	0.045**

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a three strikes law. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Table B.3: Effects of Three Strikes Laws on Violent Crime Rates (Confidence Intervals are Clustered at the Border-Region Level)

	(1)	(2)	(3)	(4)
policy	20.3 [-35.9, 76.4]	26.6 [-32.4, 88.8]	64.9 [-17.3, 149.6]	58.9 [-19.7, 131.2]
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	52	52	45	45
P-value ($H_0 : \beta = 0$)	0.519	0.455	0.130	0.140
P-value ($H_0 : \beta < 0$)	0.257	0.246	0.065*	0.085*

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Policy indicates passage of a three strikes law. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes per 100,000 people). Wild cluster bootstrap confidence intervals clustered at the border-region 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. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Table B.4: 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 Year FE	X	X	X	X
County Controls		X		X
Border-Region Weights	X	X		
Matching Weights			X	X
<i>N</i>	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. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Table B.5: 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. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Table B.6: 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
<i>N</i>	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. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Table B.7: 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
<i>N</i>	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. Coefficients show the effect of policy on the violent crime rate (i.e., number of violent crimes 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. The variables used for matching are municipal-level violent crime rates one, two, three, four, and five years prior to treatment.

Appendix C

Appendix for Chapter 3

C.1 Full Sample Summary Statistics

Additional summary statistics for my full sample are included in Table C.1.

Table C.1: Summary Statistics (Full Sample), Medical Marijuana

	Total Mean (SD)	Treated Mean (SD)	Control Mean (SD)
Violent Crime Rate	403 (475)	541 (455)	388 (475)
Male	48.7 (0.7)	49.6 (0.6)	48.7 (0.7)
Black	11.2 (7.9)	5.8 (3.8)	11.7 (8.0)
Hispanic	10.3 (9.3)	22.9 (15.7)	9.1 (7.5)
18-44 Years Old	40.2 (2.4)	40.8 (2.1)	40.1 (2.5)
No HS degree	37.3 (3.9)	40.3 (5.5)	37.0 (3.6)
In poverty	10.2 (3.6)	12.2 (3.1)	10.0 (3.6)
Unemployed	6.2 (1.6)	7.1 (1.3)	6.1 (1.6)
N	16,848	1,417	15,431

Violent Crime is a rate (number of violent crimes 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 come from Uniform Crime Reporting (UCR); all other data come from Flood et al. (2020).

BIBLIOGRAPHY

- Adda, Jérôme, Brendon McConnell, and Imran Rasul. 2014. “Crime and the depenalization of cannabis possession: Evidence from a policing experiment.” *Journal of Political Economy* 122 (5):1130–1202. [66]
- Anderson, David A. 2002. “The deterrence hypothesis and picking pockets at the pickpocket’s hanging.” *American Law and Economics Review* 4 (2):295–313. [40]
- Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist’s companion*. Princeton University Press. [29, 58]
- 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. [40]
- 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. [40]
- Banks, David L. 1988. “Histospline smoothing the Bayesian bootstrap.” *Biometrika* 75 (4):673–684. URL <http://www.jstor.org/stable/2336308>. [vii, 1, 2, 3, 4, 12, 18]
- Becker, Gary S. 1968. “Crime and punishment: An economic approach.” In *The Economic Dimensions of Crime*. Springer, 13–68. [38, 39]
- Bennett, Trevor, Katy Holloway, and David Farrington. 2008. “The statistical association between drug misuse and crime: A meta-analysis.” *Aggression and violent behavior* 13 (2):107–118. [66]
- Beran, Rudolf and Peter Hall. 1993. “Interpolated nonparametric prediction intervals and confidence intervals.” *Journal of the Royal Statistical Society: Series B*

- (*Statistical Methodology*) 55 (3):643–652. URL <http://www.jstor.org/stable/2345876>. [3]
- 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. [29, 58]
- 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. [22, 28, 52, 57]
- Chabé-Ferret, Sylvain. 2017. “Should we combine difference in differences with conditioning on pre-treatment outcomes?” Tech. rep., TSE Working Paper. [31, 60]
- Chamberlain, Gary and Guido W Imbens. 2003. “Nonparametric applications of Bayesian inference.” *Journal of Business & Economic Statistics* 21 (1):12–18. [4]
- 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. [21, 40, 41, 42]
- Cheon, Hyunjung, Scott H Decker, and Charles M Katz. 2018. “Medical marijuana and crime: Substance use and criminal behaviors in a sample of arrestees.” *Journal of Drug Issues* 48 (2):182–204. [52, 66]
- Chu, Yu-Wei Luke and Wilbur Townsend. 2019. “Joint culpability: The effects of medical marijuana laws on crime.” *Journal of Economic Behavior & Organization* 159:502–525. [51, 52]
- 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. [21, 23, 24, 25, 38, 41]
- Datta, Anusua. 2017. “California’s three strikes law revisited: Assessing the long-term

- effects of the law.” *Atlantic Economic Journal* 45 (2):225–249. [21, 41, 42]
- 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. [31, 60]
- Dragone, Davide, Giovanni Prarolo, Paolo Vanin, and Giulio Zanella. 2019. “Crime and the legalization of recreational marijuana.” *Journal of economic behavior & organization* 159:488–501. [51, 52, 68]
- 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. [22, 28, 29, 52, 57, 58]
- Durlauf, Steven N and Daniel S Nagin. 2011. “Imprisonment and crime: Can both be reduced?” *Criminology & Public Policy* 10 (1):13–54. [40]
- Efron, Bradley. 1982. *The jackknife, the bootstrap, and other resampling plans*, vol. 38. Siam. [8]
- Fisher, Ronald Aylmer. 1932. *Statistical methods for research workers*. Edinburgh: Oliver and Boyd, 4th ed. [3]
- 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>. [26, 44, 55, 70, 80, 88]
- Gavrilova, Evelina, Takuma Kamada, and Floris Zoutman. 2019. “Is legal pot crippling Mexican drug trafficking organisations? The effect of medical marijuana laws on US crime.” *The Economic Journal* 129 (617):375–407. [51, 52, 54, 62, 67]
- Goldman, Matt and David M. Kaplan. 2017. “Fractional order statistic approximation for nonparametric conditional quantile inference.” *Journal of Econometrics*

- 196 (2):331–346. URL <https://doi.org/10.1016/j.jeconom.2016.09.015>. [vii, 2, 3, 12, 13, 15, 16, 17, 18, 75]
- Goldman, Matt and David M Kaplan. 2018. “Non-parametric inference on (conditional) quantile differences and interquantile ranges, using L-statistics.” *The Econometrics Journal* 21 (2):136–169. [3]
- Goodman-Bacon, Andrew. 2018. “Difference-in-differences with variation in treatment timing.” Tech. rep., National Bureau of Economic Research. [29, 30, 54, 58]
- Hahn, Jinyong. 1997. “Bayesian bootstrap of the quantile regression estimator: A large sample study.” *International Economic Review* 38 (4):795–808. URL <http://www.jstor.org/stable/2527216>. [4]
- Helland, Eric and Alexander Tabarrok. 2007. “Does three strikes deter? A nonparametric estimation.” *Journal of Human Resources* 42 (2):309–330. [21]
- Ho, Daniel E, Kosuke Imai, Gary King, and Elizabeth A Stuart. 2007. “Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference.” *Political Analysis* 15 (3):199–236. [31, 59]
- Ho, Yvonne H. S. and Stephen M. S. Lee. 2005. “Iterated smoothed bootstrap confidence intervals for population quantiles.” *Annals of Statistics* 33 (1):437–462. URL <http://www.jstor.org/stable/3448669>. [2]
- Hofmann, Lonnie. 2020. “The effect of three strikes laws on violent crime rates.” Unpublished. [53, 55, 56, 57, 58, 59, 61, 62, 63, 64, 65]
- Huber III, Arthur, Rebecca Newman, and Daniel LaFave. 2016. “Cannabis control and crime: Medicinal use, depenalization and the war on drugs.” *The BE journal of economic analysis & policy* 16 (4). [51]
- Hunt, Priscillia, Rosalie Liccardo Pacula, and Gabriel Weinberger. 2018. “High on crime? Exploring the effects of marijuana dispensary laws on crime in California

- counties.” Tech. rep., IZA Discussion Paper. [52]
- Hutson, Alan D. 1999. “Calculating nonparametric confidence intervals for quantiles using fractional order statistics.” *Journal of Applied Statistics* 26 (3):343–353. URL <https://doi.org/10.1080/02664769922458>. [3, 17]
- 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. [21]
- Kaplan, David M. 2015. “Improved quantile inference via fixed-smoothing asymptotics and Edgeworth expansion.” *Journal of Econometrics* 185 (1):20–32. URL <https://doi.org/10.1016/j.jeconom.2014.08.011>. [3]
- 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. [22, 41, 42]
- Lo, Albert Y. 1993. “A Bayesian method for weighted sampling.” *Annals of Statistics* 21 (4):2138–2148. URL <http://www.jstor.org/stable/2242333>. [19]
- 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. [21]
- Martin, Andrew D., Kevin M. Quinn, and Jong Hee Park. 2011. “MCMCpack: Markov chain monte carlo in R.” *Journal of Statistical Software* 42 (9):22. URL <http://www.jstatsoft.org/v42/i09/>. [17]
- Marvell, Thomas B and Carlisle E Moody. 2001. “The lethal effects of three-strikes laws.” *The Journal of Legal Studies* 30 (1):89–106. [22]
- Meeden, Glen. 1993. “Noninformative nonparametric Bayesian estimation of quantiles.” *Statistics & Probability Letters* 16 (2):103–109. URL <https://doi.org/>

- 10.1016/0167-7152(93)90153-A. [2, 3, 19]
- Morris, Robert G, Michael TenEyck, James C Barnes, and Tomislav V Kovandzic. 2014. "The effect of medical marijuana laws on crime: Evidence from state panel data, 1990-2006." *PloS one* 9 (3):e92816. [51]
- 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. [31, 60]
- Neyman, Jerzy. 1937. "«Smooth test» for goodness of fit." *Skandinavisk Aktuar-tidsskrift* 20 (3-4):149-199. URL <https://doi.org/10.1080/03461238.1937.10404821>. [3]
- Paternoster, Raymond. 2010. "How much do we really know about criminal deterrence?" *The Journal of Criminal Law and Criminology* :765-824. [40]
- Pearson, Karl. 1933. "On a method of determining whether a sample of size n supposed to have been drawn From a parent population having a known probability integral has probably been drawn at random." *Biometrika* 25:379-410. URL <https://doi.org/10.1093/biomet/25.3-4.379>. [3]
- Polansky, Alan M. and William. R. Schucany. 1997. "Kernel smoothing to improve bootstrap confidence intervals." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 59 (4):821-838. URL <https://doi.org/10.1111/1467-9868.00099>. [2]
- Prison Policy Initiative. 1998. ""Three strikes" laws: Five years later." URL <https://static.prisonpolicy.org/scans/sp/3strikes.pdf>. [38, 39, 79]
- ProCon.org. 2021. "Legal Medical Marijuana States and DC." <https://medicalmarijuana.procon.org/legal-medical-marijuana-states-and-dc/>. Accessed: 2021-03-30. [53, 54, 68]

- R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. [17]
- 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. [21, 41, 42]
- 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. [29, 58]
- Roth, Jonathan. 2018. “Pre-test with caution: Event-study estimates after testing for parallel trends.” Tech. rep., Working Paper. [32, 61]
- Rubin, Donald B. 1981. “The Bayesian bootstrap.” *Annals of Statistics* 9 (1):130–134. URL <http://projecteuclid.org/euclid.aos/1176345338>. [vii, 1, 2, 3, 4, 19]
- 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. [31, 60]
- Shepard, Edward M and Paul R Blackley. 2016. “Medical marijuana and crime: Further evidence from the western states.” *Journal of Drug Issues* 46 (2):122–134. [51]
- 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. [21]
- 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. [21]

- Stuart, Elizabeth A, Gary King, Kosuke Imai, and Daniel Ho. 2011. “MatchIt: Non-parametric preprocessing for parametric causal inference.” *Journal of Statistical Software* . [31, 59, 79]
- The Federal Judicial Center. 2006. “Federal courts & what they do.” URL <https://ar.usembassy.gov/wp-content/uploads/sites/26/2016/03/FCtsWh06.pdf>. [23]
- Wilks, Samuel Stanley. 1962. *Mathematical statistics*. New York: Wiley. [5, 15]
- 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. [24]
- Worrall, John L. 2004. “The effect of three-strikes legislation on serious crime in California.” *Journal of Criminal Justice* 32 (4):283–296. [21, 41, 42]
- Wright, Valerie. 2010. *Deterrence in criminal justice: Evaluating certainty vs. severity of punishment*. Sentencing Project. [40]
- Wu, Guangzhen, Ming Wen, and Fernando A Wilson. 2021. “Impact of recreational marijuana legalization on crime: Evidence from Oregon.” *Journal of criminal justice* 72:101742. [51, 68]

VITA

Lonnie Hofmann was born in Hamilton, Ontario, Canada in 1989. He grew up in Columbia, Missouri and graduated from Rock Bridge High School in 2008. He received a Bachelor's Degree in Economics from the University of Minnesota in 2014, and a Master's Degree in Economics from the University of Missouri in 2018. He hopes to receive his Doctorate Degree in Economics from the University of Missouri in May of 2021.