CALIFORNIA'S THREE STRIKES LAW REFORMED:

What Happens to Crime Rates When Stealing a Slice of
Pizza No Longer Means Life Behind Bars

Lucy Zhu

Advisor: Professor Costas Meghir

April 21, 2021

A Senior Thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Arts in Economics

Yale University
New Haven, CT

Table of Contents

Acknowledgements **3**

Abstract **4**

1 Introduction **5**

2 Background **8**

3 Literature Review **15**

4 Theoretical Framework **18**

5 Empirical Strategy **23**

5.1 Identification Strategy *24*

5.2 Control Variables *26*

6 Data **28**

6.1 Data Source *29*

6.2 Summary Statistics *32*

7 Results **33**

7.1 Comparison of Crime Rates *33*

7.2 Testing for Differential Trends *36*

7.3 Diff-in-Diff Results *39*

7.4 Exploration of Nonlinearity *40*

7.5 Robustness Checks *44*

8 Discussion **48**

8.1 Baseline Diff-in-Diff *48*

8.2 Exploration of Nonlinearity *50*

8.3 Policy Implications *52*

9 Conclusion **55**

References **56**

Appendices **60**

Acknowledgements

I would like to thank Professor Meghir for his guidance and mentorship, and for pushing me to think more deeply and independently. I would also like to thank Vitor Possebom for the many engaging conversations about crime and economics that inspired me each week to keep working on this essay. I could not have finished this project without the support of my friends whose encouragement and care kept me grounded and connected. In particular, I'm extremely grateful to Dave Jiang for motivating me with his wit and intelligence when things got tough, and to Amber Hu for the regular conversation and laughter. All errors are my own.

Abstract

California's Three Strikes Law (TSL) was the harshest habitual offender law in the nation. It provided for mandatory life imprisonment for all repeat offenders, regardless of the crime committed. In 2012, California voters passed Proposition 36 (Prop. 36), which reformed the TSL to eliminate mandatory life sentences for non-violent offenders. This paper examines the effect of Prop. 36 on crime rates in order to determine how it compares with the TSL as a sentencing policy. By employing a difference-in-differences approach at the county-level and using 2000-2017 crime data from the FBI's Uniform Crime Reporting Program, this study finds that Prop. 36 leads to a significant decrease in the violent crime rate while it had no effect on the property crime rate. This study also finds that the TSL had failed to reduce crime.

1 Introduction

In the summer of 1994, 27-year-old Jerry Dewayne Williams stole a slice of pepperoni pizza from a group of children near Redondo Beach. Six months later, he was convicted of felony petty theft and sentenced to life in prison under California's Three Strikes Law (TSL).[1] Petty theft is typically charged as a misdemeanor, but because of Williams' earlier convictions, prosecutors decided to file it as a felony instead. Under the TSL, if an offender has two prior convictions for a serious or violent offense, she will receive a mandatory life sentence for any subsequent felony that she commits, regardless of the gravity of the crime. Since Williams had prior convictions for robbery and attempted robbery, both of which are considered "serious" offenses under California's Penal Code, his subsequent conviction of felony petty theft mandated that he would be sentenced to life imprisonment. As a result, for stealing a slice of pizza, Williams received the same sentence as if he had committed a murder or molested a child.

The TSL was enacted in 1994 in response to the dramatic rise in California's violent crime rate in the 1980s.[2] The TSL was designed to target repeat offenders since it was assumed that they were responsible for the rise in violent crime. Legislators promised it would keep dangerous criminals off the streets and reduce violent crime. However, a decade after the TSL's enactment, the majority of those locked up under the TSL were non-violent offenders. Furthermore, because the TSL provided for substantial sentence enhancements, especially for non-violent offenders sentenced to life in prison, it contributed to California's exploding prison population. By 2004, strikers sentenced under the TSL constituted over 25% of the prison population.[3]

The passage of the TSL was as much a political instrument as it was a solution to the rising crime rates. 1994 was an election year, and incumbent Republican Pete Wilson was running for re-election in a competitive gubernatorial race. Although the violent crime rate in California began declining in 1993, as shown in Figure 1 below, Wilson centered his campaign around crime in hopes of exploiting the public's prejudicial attitude towards criminals. He was an outspoken advocate of the TSL, and although some legislators questioned its efficacy and excessive

---

[1] The Associated Press, "25 Years for a Slice of Pizza," *The New York Times*, March 5, 1995, sec. U.S., https://www.nytimes.com/1995/03/05/us/25-years-for-a-slice-of-pizza.html.

[2] Radha Iyengar, "I'd Rather Be Hanged for a Sheep than a Lamb: The Unintended Consequences of 'Three-Strikes' Laws" (National Bureau of Economic Research, February 7, 2008), https://doi.org/10.3386/w13784.

[3] Brian Brown and Greg Jolivette, "A Primer: Three Strikes - The Impact After More Than a Decade," Legislative Analyst's Office, October 2005, https://lao.ca.gov/2005/3_strikes/3_strikes_102005.htm.

penalties, very few legislators were willing to officially oppose it – their own political careers were on the line. Legislators feared that resisting the TSL or even proposing a more modest version would make them appear soft on crime.[4] In 1994, the TSL was passed by the legislature.

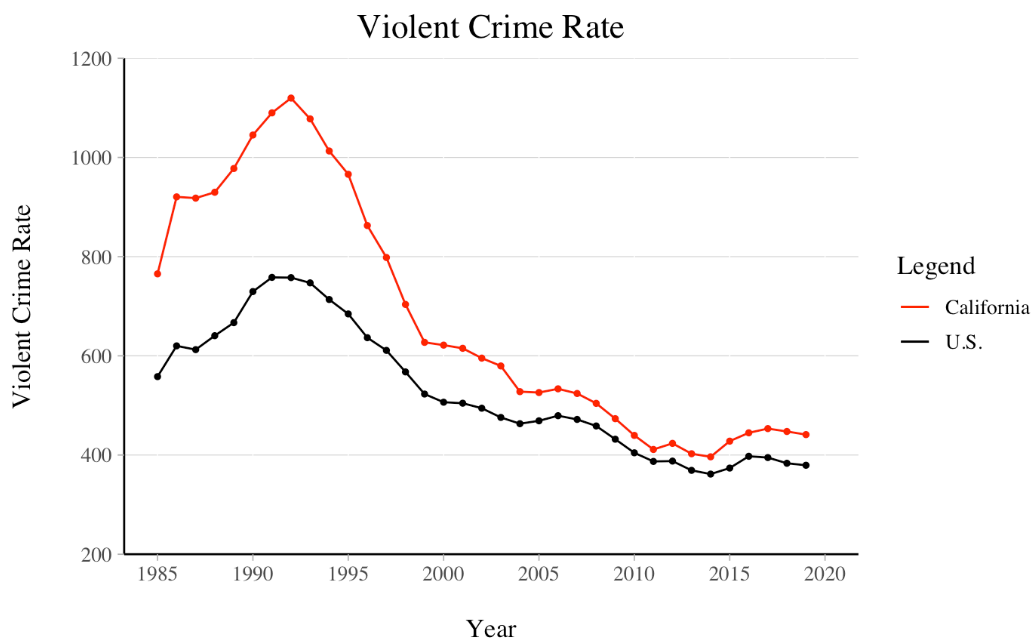


Figure 1: The violent crime rate in California and the United States from 1985-2019. Data is from the FBI's Uniform Crime Reporting Program

Ten years later, efforts to eliminate mandatory life sentences for non-violent offenders culminated in Proposition 66, which appeared on the November 2004 California ballot. Polls indicated that there was overwhelming support for the initiative by voters. However, just days before the election, then-Governor Arnold Schwarzenegger launched a series of graphic advertising campaigns against the proposition. The ads portrayed non-violent offenders as hardened criminals and warned that Proposition 66 would release 26,000 of them back onto the streets.[5] The goal was to scare the public, and it worked. Proposition 66 suffered a narrow defeat.[6]

---

[4] Michael Vitiello, "Three Strikes: Can We Return to Rationality?," *The Journal of Criminal Law and Criminology (1973-)* 87, no. 2 (1997): 395, https://doi.org/10.2307/1143951.

[5] The Associated Press, "11th-Hour Attempt Overturns Prop 66," The Santa Clara, November 3, 2004, https://www.thesantaclara.org/blog/11th-hour-attempt-overturns-prop-66.

[6] California voters rejected Proposition 66, with 53% voting against it.

In 2006, California's prison population reached its peak of about 173,000 inmates.[7] Meanwhile, California's violent crime rate was at a historic low, reaching levels not seen since the early 1970s.[8] Public discourse began to turn towards the injustices in the criminal justice system and the social implications of mass incarceration. In 2011, the U.S. Supreme Court ruled in *Brown v. Plata* (2011) that California prison conditions were unconstitutional and ordered the state to reduce its prison population. As a result, California pursued a series of criminal justice reforms to alleviate its overcrowded prisons.[9] One of them was the Three Strikes Reform Act (Proposition 36), which tried once again to eliminate mandatory life sentences for non-violent offenders from the TSL. In 2012, California voters approved the ballot measure, nearly two decades after the TSL was passed into law.

Several other criminal justice reforms were passed after Proposition 36 (Prop. 36). However, in recent years, police unions and opponents of criminal justice reform have increasingly called for measures to roll back these changes. Opponents cite the slight uptick in violent crime in 2015 and 2016 as evidence that the changes have gone too far.[10] In 2017, Assemblyman Jim Cooper (D-Elk Grove) announced that he was working on a ballot measure that would reverse the recent criminal justice reforms. He declared that harsher "consequences" are necessary because non-violent offenders are usually "linked to more serious violent crimes of rape and murder" in the future.[11]

However, there has been no empirical evidence that supports these claims. No studies have shown that reducing the TSL's excessively harsh punishments on non-violent offenders led to more violence. The motivation for this paper is thus to examine whether Prop. 36 actually increased the violent crime rate by eliminating life sentences for non-violent offenders. By examining this question, we can ensure that the harsh punishments provided by the TSL, which affect the lives of thousands of people, are backed by empirical evidence and not enacted for the sake of furthering someone's political agenda or maintained through the use of fear-based marketing tactics.

---

[7] Ashley Gabbard et al., "Offender Data Points," 2019, 177.

[8] Magnus Lofstrom and Brandon Martin, "Crime Trends in California," *Public Policy Institute of California* (blog), February 2021, https://www.ppic.org/publication/crime-trends-in-california/.

[9] Robert Weisberg, "The Wild West of Sentencing Reform: Lessons from California," *Crime and Justice* 48 (February 8, 2019): 35–77, https://doi.org/10.1086/701714.

[10] Tim Arango, "In California, Criminal Justice Reform Offers a Lesson for the Nation (Published 2019)," *The New York Times*, January 21, 2019, sec. U.S., https://www.nytimes.com/2019/01/21/us/california-incarceration-reduction-penalties.html.

[11] Jessica Pishko, "In Liberal California, A Crusader Against Criminal Justice Reform," The Appeal, August 19, 2019, https://theappeal.org/in-liberal-california-a-crusader-against-criminal-justice-reform/.

In this paper, we will examine the effect of Prop. 36 on the violent crime rate, the property crime rate, and the overall crime rate, relative to the TSL. We will employ a difference-in-differences approach with 2000-2017 crime data from the FBI's Uniform Crime Reporting Program (UCR). In order to identify the effect, we exploit the variation in how strictly each county implemented the TSL. Additionally, we also examine whether the effect of Prop. 36 depends on each county's strictness level.

We find that Prop. 36 decreased the violent crime rate by 11.5% in counties that closely adhered to the TSL but did not affect the violent crime rate in other counties. Our results suggest that Prop. 36 is more effective than the TSL in reducing violent crime while it is just as effective as the more lenient versions of the TSL that some counties had been using. We also find that Prop. 36 had no significant effect on either the property crime rate or the overall crime rate.

The remainder of the paper is structured as follows. Section 2 presents the mechanics of the TSL and Prop. 36 as well as their legislative history. Section 3 discusses prior studies on the effectiveness of the TSL. Section 4 provides an enhanced model of crime for understanding criminal behavior under the TSL. Section 5 specifies the difference-in-differences model and describes the identification strategy. Section 6 provides the data sources, and Section 7 presents the results. Section 8 discusses the results and considers policy implications, and Section 9 concludes.

2 Background

In 1992, Joe Davis attempted to rob 18-year-old Kimber Reynolds outside a restaurant in Fresno, California. When Reynolds refused to hand over her purse, Davis shot her in the head. 26 hours later, she died. It was later discovered that Davis had been released only two months earlier from prison after serving out a sentence for auto theft. Reynolds' death sparked public outrage and prompted her father, Mike Reynolds, to take action. What resulted was the draft of what is now known as the Three Strikes Law.[12] Crime in California, especially violent crime, had shot up to alarming levels over the past decade. The public feared for its safety and was eager to support a new sentencing law that promised to reduce violent crime. In 1994, the Three Strikes Law was passed both by the legislature and by ballot initiative, with 72% of voters in favor of the new measure.[13]

---

[12] George Skelton, "A Father's Crusade Born From Pain," Los Angeles Times, December 9, 1993, https://www.latimes.com/archives/la-xpm-1993-12-09-mn-65402-story.html.

[13] "California Proposition 184, Three Strikes Sentencing Initiative (1994) - Ballotpedia," accessed April 19, 2021, https://ballotpedia.org/California_Proposition_184,_Three_Strikes_Sentencing_Initiative_(1994).

The Three Strikes Law (TSL) aims to reduce violent crime by targeting repeat offenders. It rests on the assumption that repeat offenders are responsible for a large percentage of violent crimes. There are two components to the TSL – the two-strikes provision and the three-strikes provision.[14] Under the two-strikes provision, if a person has one previous conviction for a serious or violent felony ("strikeable offense"), a subsequent offense of *any* felony will lead to a mandatory sentence of twice the term otherwise provided for the current conviction. Similarly, under the three-strikes provision, if a person has two or more previous convictions for a strikeable offense, *any* felony committed subsequently will trigger a mandatory sentence of 25 years to life in prison.[15] Table 1 below lists the serious and violent crimes that constitute strikeable offenses.

[14] Cal. Penal Code § 667(b)-(i) and 1170.12

[15] The minimum term for third-strikers is calculated as the greater of: (1) three times the term otherwise provided for the current felony conviction; (2) 25 years; or (3) the term provided by law for the current conviction, including any applicable sentence enhancements.

Table 1: Strikeable offenses

Type of Crime	Specific Offense
Homicide	Murder
Sexual offenses	Rape
	Lewd act on child
	Continual sexual abuse of child
	Penetration by foreign object
	Sexual penetration by force
	Sodomy by force
	oral copulation by force
Robbery	Robbery
Felony assault	Attempted murder
	Assault with a deadly weapon on a peace officer
	Assault with a deadly weapon by an inmate
	Assault with intent to rape or rob
Other crimes against persons	Any felony resulting in bodily harm
	Arson causing bodily harm
	Carjacking
	Exploding device with intent to injure
	Exploding device with intent to murder
	Kidnapping
	Mayhem
Property crimes	Arson
	Burglary of occupied dwelling
	Grand theft with firearm
Drug offenses	Drug sales to minors
Weapons offense	Any felony with deadly weapon
	Any felony where firearm is used

Source: John Clark, James Austin, and Alan Henry, "Three Strikes and You're Out: A Review of State Legislation," 1997.

The TSL was built on the theory of deterrence and incapacitation. Deterrence is the philosophy that the threat of punishment will prevent an individual from committing a crime if it outweighs the benefits of her doing so. It rests on the assumption that an individual is capable of perceiving the certainty and severity of punishment and that she is a rational actor.[16, 17] However, lawmakers behind the TSL concluded that regarding career criminals, the certainty of punishment is more

[16] Kelli D. Tomlinson, "An Examination of Deterrence Theory: Where Do We Stand," *Federal Probation* 80, no. 3 (2016): 33–38.

[17] Certainty of punishment relates to the likelihood of being apprehended as well as the likelihood of conviction if apprehended. For example, more judicial discretion means less certainty of punishment.

important for deterring them than the severity of punishment. They argued that since these criminals have already shown through their pattern of illegal activities that they do not care about the specific length of incarceration, the only "deterrent value" of incarceration as a punishment is its existence as an inescapable consequence. As a result, the TSL used lengthy mandatory sentences as a way to communicate to repeat offenders the "reality and inevitability of punishment."[18]

The TSL also seeks to reduce crime through incapacitation, where offenders are removed from society to prevent future criminal behavior. The three-strikes provision is the most prominent example of this approach – eligible repeat offenders are locked away and isolated from society for at least 25 years. The assumptions underlying incapacitation as an effective method for reducing crime are that (1) the courts can accurately identify career criminals and (2) these criminals will continue to offend if they are not locked away.[19]

Although the TSL was intended to limit discretion in the criminal justice system as a mandatory sentencing law, there was a great deal of variation among California counties in terms of implementation. The TSL allows for prosecutors to dismiss prior strike convictions "in furtherance of justice" for sentencing purposes.[20] For example, prosecutors in Los Angeles were directed by their District Attorney to only apply the TSL to third-strike cases if the offense is considered "serious" under the Penal Code. Meanwhile, prosecutors in Kern County were told to give out life sentences to all eligible offenders. Furthermore, in 1996, the California Supreme Court ruled in *People v. Romero* (Cal. 1996) that trial judges also have the authority to dismiss prior strikes "in furtherance of justice."[21] Thus, both prosecutors and judges hold discretionary powers under the TSL when sentencing repeat offenders. Discretion can also be exercised when dealing with certain offenses known as "wobblers."[22] In such cases, prosecutors can decide whether to charge the wobbler as a felony or a misdemeanor.[23] If a wobbler is charged as a

---

[18] James A. Ardaiz, "California's Three Strikes Law: History, Expectations, Consequences," *McGeorge Law Review* 32 (2000), https://heinonline.org/HOL/Page?handle=hein.journals/mcglr32&id=11&div=&collection=.

[19] James Austin et al., "Three Strikes and You're Out: The Implementation and Impart of Strike Laws," March 6, 2000.

[20] Michael Romano, "Divining the Spirit of California's Three Strikes Law," *Federal Sentencing Reporter* 22, no. 3 (2010): 171–75, https://doi.org/10.1525/fsr.2010.22.3.171.

[21] People v. Superior Court (Romero), 13 Cal. 4th 497, 504 (1996).

[22] Common wobblers in California include forgery, grand theft, receiving stolen property, and petty theft with a prior.

[23] Elsa Chen, "The Liberation Hypothesis and Racial and Ethnic Disparities in the Application of California's Three Strikes Law," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 2008), https://papers.ssrn.com/abstract=2189506.

felony, then it is eligible to receive a strike sentence under the TSL. Otherwise, the TSL will not apply.[24]

Soon after its passage, California's TSL became a source of immense controversy due to its tough sentencing provisions. Although 24 states had passed their own version of the three-strikes law by 1995, California's version remained the most severe in the nation.[25] Three factors contributed to its severity: (1) the law's all-inclusive strike-out zone; (2) the existence of a two-strikes provision in addition to a three-strikes provision; and (3) the aggressiveness with which California applied the law.

No other state implemented an all-inclusive strike-out zone, where any felony could trigger a life sentence on the third strike.[26] This meant that in addition to strikeable offenses, repeat offenders in California could also "strike-out" on non-serious or non-violent offenses ("non-strikeable offenses"), such as drug possession or petty theft. The rationale behind this tough provision rests on the assumption that those who have committed multiple strikeable offenses have a disposition toward violent behavior and will continue to pose a threat to society if they are not punished. Thus, as argued by the lawmakers behind the TSL, it does not matter the type of felony that a repeat offender commits on her third strike since she is likely to offend again in the future. Even if she commits a low-level felony, a life sentence is necessary to remove her from society before she commits a more violent offense down the line.[27] No other state imposes a life sentence for a non-violent offense, even if it is committed by a repeat offender.[28] Critics of the TSL have argued that the three-strikes provision constitutes "cruel and unusual punishment" under the Eighth Amendment, but the U.S. Supreme Court ruled that it was justified in the interest of "public safety" and ultimately upheld the constitutionality of the law in *Ewing v. California* (2003).[29]

---

[24] Judges can also decide whether to sentence the wobbler as a felony or a misdemeanor. The former can trigger a longer sentence under the TSL while the latter will not.

[25] "Severity" as defined by the number of offenders sentenced under the TSL.

[26] Austin et al., ibid.

[27] Ardaiz, ibid.

[28] Romano, ibid.

[29] Ewing v. California, 538 U.S. 11 (2003)

Following its passage, the TSL dramatically increased California's prison inmate population. By the end of 2004—a decade after the TSL's enactment—second- and third-strikers made up over 25% of the prison population (see Figure 2 below). The majority of this increase was the result of the two-strikes provision: out of the almost 43,000 strikers in prison in December 2004, over 80% of them were second-strikers.[30] Although a few other states with three-strikes laws also had a two-strikes provision, they applied it to a much shorter list of offenses. For example, Georgia has a relatively severe two-strikes provision where second-strikers are automatically sentenced to life in prison. However, this sentence can only be triggered by a narrow set of violent offenses, including murder, rape, and aggravated child molestation.[31] Meanwhile, California's two-strikes provision, like its three-strikes provision, can be triggered by *any* felony. Furthermore, California applied the TSL more frequently than any other state with a three-strikes law. For example, Washington—the first state to pass a three-strikes law in 1993—had only sentenced 115 offenders under its new law by 1998. California, on the other hand, had sentenced nearly 40,000 offenders under the TSL by then. As a result, given the broad scope of the TSL and the aggressiveness with which it was applied, California saw

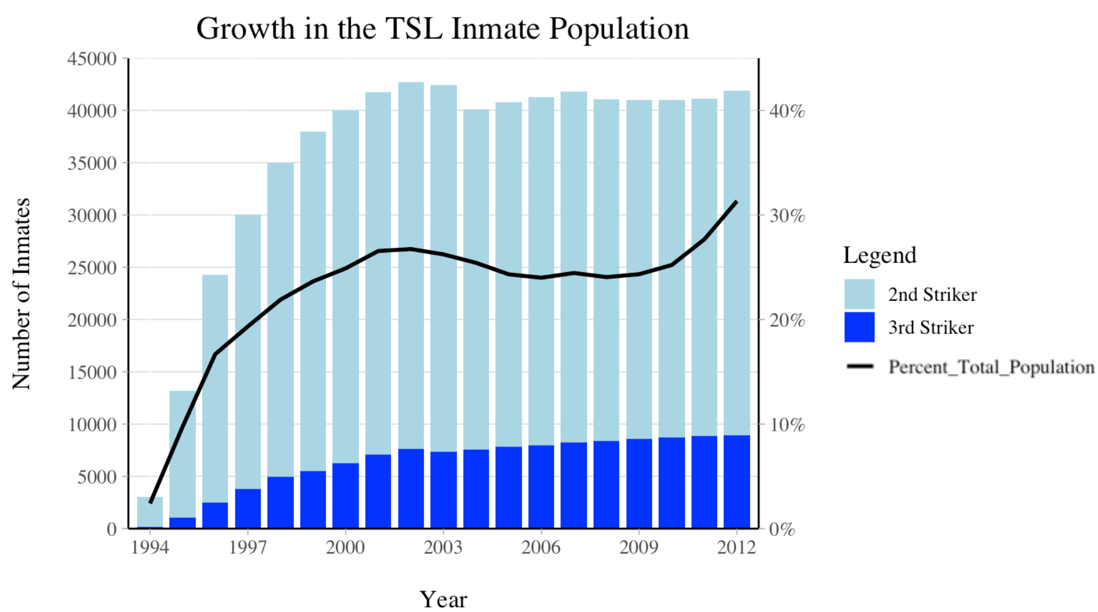


Figure 2: The growth in the number of second- and third-strikers among the prison population 1994-2012. The y-axis on the left represents the number of striker inmates. The y-axis on the right represents strikers as percent of total prison population. Data is from California Department of Corrections and Rehabilitation (CDCR).

[30] Brown and Jolivette, ibid.

[31] Austin et al., ibid.

its prison population explode to about 200% of design capacity by 2000, if not earlier.[32]

As California's incarceration rate skyrocketed, public sentiment shifted towards criminal justice reform. Mass incarceration and unsafe prison conditions made its way to the forefront of public discourse. In 2009, a federal court ordered California to reduce its prison population to 137.5% of total design capacity, but California appealed the decision. In 2011, the Supreme Court ruled that California's overcrowded prisons were a violation of the 8th Amendment and upheld the federal court's order.[33] The Supreme Court's mandate, along with popular anti-incarceration campaigns, led to a series of criminal justice reforms, including Prop. 36.[34]

Prop. 36 modified the TSL to "restore the original intent" of the TSL such that life sentences are only used to punish dangerous criminals.[35] Under Prop. 36, a mandatory life sentence can only be imposed when a strikeable offense is committed.[36] However, if a third-striker has any prior convictions for rape, murder, or child molestation, she will continue to be punished with life imprisonment regardless of her current conviction. Prop. 36 also provided for the potential re-sentencing of current third-strike inmates who otherwise would not have received a life sentence under Prop. 36 (i.e. certain non-violent third-strike inmates). Re-sentencing would only occur if a judge determines that doing so will not create "unreasonable risk to public safety."[37]

In November 2012, Prop. 36 successfully passed by ballot initiative with 69% of voters in favor of the reform. Eighteen months after its passage, Prop. 36 had released over 1,600 inmates from prison with its re-sentencing provision. The recidivism rate of these released prisoners was found to be 1.3%, which is much lower than that of other inmates released during the same time period (over 30%).[38,]

---

32 Brown v. Plata, 563 U.S. 493 (2011)

33 Ibid.

34 Weisberg, ibid.

35 "Text of California Proposition 36 (November 2012)," Ballotpedia, accessed December 21, 2020, https://ballotpedia.org/Text_of_California_Proposition_36_(November_2012).

36 Certain non-strikeable offenses that involve sex, drugs, or firearm possession will still trigger a life sentence under Prop. 36.

37 "California Proposition 36, Changes to Three Strikes Sentencing Initiative (2012)," Ballotpedia, accessed April 19, 2021, https://ballotpedia.org/California_Proposition_36,_Changes_to_Three_Strikes_Sentencing_Initiative_(2012).

38 "Recidivism" is defined as being released from custody and then getting re-incarcerated in state prison for a new crime.

[39] Since the introduction of Prop. 36, no other major changes have been made to the TSL.

3 Literature Review

There have been multiple studies on the effects of the TSL on crime, but the conclusions have been mixed. Some have found a strong deterrent effect while others have found no effect at all. Some have even found a positive effect. One econometric challenge that these studies face when trying to identify the effect of the TSL is separating the pure deterrent effect from the incapacitation effect. The latter arises due to the sentence enhancements provided by the TSL, as described in Section 2 above. Distinguishing between the two effects is important because it is more desirable from a policy perspective to reduce crime through deterrence than incapacitation. The latter is much more expensive to achieve since it requires the state to provide and to operate the necessary prison facilities. Additionally, incapacitation also requires the state to bear the cost of caring for an aging inmate population.

In an earlier study, Marvell and Moody (2001) examined the effects of three-strikes laws on crime rates in all 50 states.[40] They used a fixed-effects model with state-level data from 1970-1998, and a dummy variable was created to indicate the presence of a three-strikes law. They modeled the incapacitation effect of the law with a linear trend variable that started in the year after each law was passed. They found that three-strikes laws actually increase the number of homicides by 10-12% in the short run and 23-29% in the long run. They attributed this increase to third-strikers killing victims and witnesses in order to avoid the harsher punishments provided by these laws. They also found that three-strikes laws have no deterrent or incapacitation effects.

Shepherd (2002) improved upon Marvell and Moody's (2001) analysis by using county-level data instead. She argued that a state-level study of the three-strikes law does not capture the variation in how counties apply the law.[41] According to Shepherd, failing to account for this variation would likely lead to an underestimation of the law's effectiveness. Furthermore, not only is there variation in enforcement at the county level but states with three-strikes laws also vary in how

[39] Stanford Law School Three Strikes Project and NAACP Legal Defense and Education Fund, "Proposition 36 Progress Report: Over 1,500 Prisoners Released; Historically Low Recidivism Rate," April 2014, https://law.stanford.edu/publications/proposition-36-progress-report-over-1500-prisoners-released-historically-low-recidivism-rate/.

[40] Thomas B. Marvell and Carlisle E. Moody, "The Lethal Effects of Three-Strikes Laws," *The Journal of Legal Studies* 30, no. 1 (January 1, 2001): 89–106, https://doi.org/10.1086/468112.

[41] Joanna M. Shepherd, "Fear of the First Strike: The Full Deterrent Effect of California's Two- and Three-Strikes Legislation," *The Journal of Legal Studies* 31, no. 1 (January 1, 2002): 159–201, https://doi.org/10.1086/324660.

frequently they apply them. By classifying states that regularly apply their laws in the same group as states that do not, Shepherd argued that Marvell and Moody further underestimate the effects of the law. Thus, as Shepherd argued, studies on the three-strikes law should only focus on states that regularly enforce the law.

Using panel data on all 58 California counties from 1983-1996, Shepherd employed a two-stage least squares regression to examine the deterrent effect of the TSL on crime rates in California. In order to identify the deterrent effect of the TSL, Shepherd exploited the variation among counties in how aggressively they applied the law.[42] She argued that she was capturing the pure deterrent effect since she was looking at the effect of the TSL immediately after its passage, before the incapacitation effect had a chance to affect the data. She found that the TSL had a "full deterrent" effect in that it not only deterred those who were facing their last strike, but it also prevented individuals from committing their first strike. Shepherd hypothesized that this was due to individuals "fearing" a first strike since it would bring them one strike closer to the harsher penalties provided by the TSL.

Although Shepherd improved upon previous research by accounting for the county variation in TSL enforcement, her choice of instrumental variables violates the exclusion restriction. For example, she uses expenditure on the police in her model to instrument for the probability of arrest. In order for police expenditure to be a valid instrument, it must be uncorrelated with her outcome variable (crime rates). However, previous studies have shown that crime and police expenditure are closely related through a simultaneous relationship.[43] More crime leads to more spending on law enforcement, which in turn leads to fewer crimes. Thus, in Shepherd's case, using police expenditure as an instrumental variable is invalid.

Worrall (2004) used a similar identification strategy as Shepherd but obtained a different result. He found that after controlling for county-specific time trends, the TSL had no deterrent or incapacitation effect on crime.[44] Worrall used county-level data from 1989-2000 and ran a fixed effects regression model. He estimated the incapacitation effect by using the same linear trend method as Marvell

---

[42] Shepherd defines aggressiveness for each county as the probability of receiving a strike sentence, which is calculated by dividing the number of offenders receiving a strike sentence by the number of offenders admitted to prison.

[43] Lee R. McPheters and William B. Stronge, "Law Enforcement Expenditures and Urban Crime," *National Tax Journal* 27, no. 4 (December 1, 1974): 633–44, https://doi.org/10.1086/NTJ41861994; Pamela Irving Jackson and Leo Carroll, "Race and the War on Crime: The Sociopolitical Determinants of Municipal Police Expenditures in 90 Non-Southern U.S. Cities," *American Sociological Review* 46, no. 3 (1981): 290–305, https://doi.org/10.2307/2095061.

[44] John L. Worrall, "The Effect of Three-Strikes Legislation on Serious Crime in California," *Journal of Criminal Justice* 32, no. 4 (July 1, 2004): 283–96, https://doi.org/10.1016/j.jcrimjus.2004.04.001.

and Moody (2001). But unlike previous research, he also controlled for spatial autocorrelation in crime by adding a variable that represented the average crime rate among contiguous counties. Worrall cited this spatial autocorrelation variable as a potential reason for why his results differed from that of Shepherd and other papers.

In another study of the TSL, Helland and Tabarrok (2007) found a strong deterrent effect using recidivism data on a random sample of California prisoners who were released in 1994.[45] They found that among repeat offenders with two prior strikes, the TSL reduced their felony arrest rates by 17-20%.[46] Helland and Tabarrok identified this effect by using a nonparametric Kaplan-Meier model to compare the subsequent arrest profiles of those who had been convicted of two strikeable offenses with those who had went to trial for two strikeable offenses but had only received a conviction for one of them. In doing so, they assumed that the number of strike convictions handed down to each offender was exogenous, i.e. an offender's trial outcome had nothing to do with her propensity for crime. However, this assumption regarding the exogeneity of trial outcomes is problematic given how each county applies the TSL differently. For example, as discussed in Section 2, judges are allowed under *People v. Romero* to dismiss prior strikes as they see fit. Thus, Helland and Tabarrok's estimation of the effect of the TSL is likely to be biased by the discretion that exists in the criminal justice system.

Iyengar (2008) examined the effect of the TSL by comparing the criminal propensity of individuals with similar criminal histories but different strike eligibility before and after the enactment of the TSL.[47] She used offender-level records for a sample of individuals who were arrested in California from 1990-1999. She found that the TSL reduced criminal activity among second-strike eligible offenders by 20% and third-strike eligible offenders by 28%. She also identified two unintended consequences of the TSL. First, offenders were more likely to commit violent crimes due to the "flattening of the penalty gradient."[48] Second, the TSL had a migration effect where it increased the number of offenders who migrated

---

[45] Eric Helland and Alexander Tabarrok, "Does Three Strikes Deter? A Nonparametric Estimation," *Journal of Human Resources* XLII, no. 2 (March 31, 2007): 309–30, https://doi.org/10.3368/jhr.XLII.2.309.

[46] Unlike Shepherd's analysis, which was able to measure the full deterrent effect of the TSL, Helland and Tabarrok's method only captures its marginal deterrent effect.

[47] Iyengar, ibid.

[48] "Flattening of the penalty gradient" refers to how the TSL reduces the difference in severity between the penalty for violent crimes and that of non-violent crimes. For example, third-strikers under the TSL receive the same penalty regardless of the crime that they commit whereas before the TSL, those who commit a violent crime get a more severe punishment than if they had committed a non-violent crime.

from California to commit crimes in neighboring states since their laws were much more lenient than the TSL.[49]

Although there have been many studies on the efficacy of the TSL in general, little is known on whether the law had been too severe. In other words, although some studies have found that the TSL had a deterrent effect, it is unclear what provision of the TSL accounts for such an effect. For example, Iyengar (2008) found that the TSL reduced criminal activity among third-strike eligible offenders, but from her research, we do not know whether this reduction was due, specifically, to the extremely harsh punishment that the TSL imposed on non-violent third-strikers. This is an important nuance to make because it provides insight on the necessity of the TSL's harshest aspect as a sentencing law. If criminal activity was reduced as a result of imposing life sentences on non-violent offenders, then such a severe punishment may be justified. However, if this is not the case, then sentencing non-violent offenders to life in prison was unnecessary, and the TSL had been excessively harsh.

In this paper, we attempt to fill this gap in the TSL literature by utilizing the natural experiment offered by the introduction of Prop. 36. If sentencing non-violent offenders to life in prison had led to a reduction in crime under the TSL, then we should expect to see an increase in crime after Prop. 36. Although Prop. 36 modified the TSL nearly a decade ago, no studies have been done so far on its efficacy compared to the original, more severe TSL.

4 Theoretical Framework

The economic model of crime rests on the assumption that individuals respond to incentives. In a simple binary model, where legitimate and illegitimate activities are mutually exclusive, an individual will choose to allocate her time and resources based on the expected utility she anticipates from each activity. She will commit an offense only if her expected utility from illegitimate activities (i.e. crime) exceeds the utility that she gets from legitimate alternatives.[50] The expected utility from committing an offense is given by the function

$$EU_j = p_j U_j(Y_{ji} - f_j) + (1 - p_j) U_j(Y_{ji}) \tag{1}$$

[49] Specifically, these offenders include those who are eligible for a second- or third-strike sentence under the TSL if they commit another crime.

[50] Gary S. Becker, "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76, no. 2 (1968): 169–217.

where p_j is the individual's probability of apprehension and conviction per offense; U_j is her utility function; Y_{ji} is the monetary and psychic income that she receives per offense; and f_j is her punishment per offense.

Although the standard model of crime captures the immediate losses incurred if the offender is convicted for the current offense, it does not capture the additional costs inherent to the penalty structure of the TSL. Specifically, it does not reflect the increased risk of receiving a more severe punishment in the future if the current conviction is a strikeable offense. Shepherd (2002) addresses this shortfall by enhancing the standard crime model to include the deterrent effect of delayed punishment. This paper will use Shepherd's model as a framework in its analysis of Prop. 36 and crime.

Suppose an offender gets convicted of a strikeable offense and receives her first strike. She will suffer immediate losses associated with conviction, such as the confiscation of her loot or the loss of earnings due to imprisonment. She will also suffer non-monetary losses, such as the psychic cost of imprisonment. Both monetary and non-monetary losses are represented by f_j in the expected utility function. In addition to incurring f_j, she will also be one strike closer to receiving a second- or third-strike conviction in the future. In other words, the offender has lost the option of committing another first-strike, which will be represented by the variable δ_j. This loss is costly to the offender since the punishment for second- and third-strikers is more severe than what they would have received without the TSL. As a result, when deciding whether to commit an illegal act under the TSL framework, an offender must also consider the possibility of receiving higher penalties in the future (delayed punishment). The new expected utility function is thus given by

$$EU_j = p_j U_j(Y_{ji} - f_j - \delta_j) + (1 - p_j)U_j(Y_{ji}) \qquad (2)$$

where $\delta_j > 0$, which represents the increased risk of delayed punishment if convicted of a strikeable offense.[51] We can either assume that δ_j increases with the number of strikes due to increasingly harsh punishments or that it does not change; regardless, the model will give the same conclusions.

Assuming that the utility function of the offender is strictly increasing, the model as presented implies that her expected utility from offending is decreasing with higher p_j, f_j, and δ_j. This suggests that potential offenders are deterred when

[51] A non-strikeable offense will not result in harsher punishment in the future ($\delta_j = 0$). This is because a non-strikeable offense committed with no prior strikes will not trigger strike sentencing. If an offender has one prior strike, a current conviction for a non-strikeable offense will lead to harsher sentencing in the current period under the TSL's two-strikes provision, but it won't count as a second-strike. In other words, the offender is not any closer to receiving a third-strike conviction than she was before.

there is a higher chance of apprehension and conviction, a more severe punishment associated with the offense, or a greater risk of harsher punishment in the future. This model is consistent with deterrence theory, which says that deterrence occurs when an individual perceives greater certainty and severity of punishment.[52] We can see this from our model by taking the derivative of equation (2) with respect to p_j, f_j, and δ_j:

$$\frac{d(EU_j)}{dp_j} = U_j(Y_{ji} - f_j - \delta_j) - U_j(Y_{ji}) < 0 \tag{3}$$

$$\frac{d(EU_j)}{df_j} = -p_j U_j'(Y_{ji} - f_j - \delta_j) < 0 \tag{4}$$

$$\frac{d(EU_j)}{d\delta_j} = -p_j U_j'(Y_{ji} - f_j - \delta_j) < 0 \tag{5}$$

The variable f_j, which represents the direct punishment per offense, can take on one of two values: F_{TSL} or F_0, where $F_{TSL} > F_0$. The value F_{TSL} represents the harsher penalty that the TSL imposes on second- and third-strikers. This includes lost earnings and non-monetary damages associated with a longer prison sentence. As a result, f_j equals F_{TSL} when the offense committed results in a second- or third-strike conviction. On the other hand, F_0 represents the punishment that's typically given for a particular offense. Thus, f_j equals F_0 if either the offender has no previous strike convictions or if the offense does not count as a strike under the TSL.

Because counties varied in how they enforced the TSL, the magnitude of δ_j depends on the county in which the strikeable offense is committed. As discussed earlier, strict counties are more likely than lenient counties to sentence offenders under the TSL. This means that an offender who commits a strikeable offense in a strict county will face a higher risk of more severe punishment. Thus, the value of δ_j increases with the strictness of a county. To simplify the model for analysis, we will consider the binary case where a county is either strict or lenient:

$$EU_{Sj} = p_j U_j(Y_{ji} - f_{Sj} - \delta_{Sj}) + (1 - p_j) U_j(Y_{ji}) \tag{6}$$

$$EU_{Lj} = p_j U_j(Y_{ji} - f_{Lj} - \delta_{Lj}) + (1 - p_j) U_j(Y_{ji}) \tag{7}$$

where $\delta_{Sj} > \delta_{Lj}$. EU_{Sj} and EU_{Lj} are the expected utility of an individual from a strict and lenient county, respectively. Equations (6) and (7) show that, holding all else

---

[52] Tomlinson, ibid.

constant, it is more costly to commit a strikeable offense in a strict county under the TSL.

Now, we will consider how the expected utility of an offender changes after the introduction of Prop. 36. It is important to note first that the model assumes that the offender is well-informed about Prop. 36 and its consequences. This assumption implies that, given that the offender is rational, the offender is able to assess the true costs and benefits of committing a crime.

After Prop. 36 modified the TSL, the offenders most affected by the reform were those from strict counties who were choosing whether to commit a strikeable offense.[53] Since only a serious or violent offense can now trigger a life sentence on the third strike, Prop. 36 effectively reduced the risk of receiving a harsher penalty in the future. For example, even though an offender who gets convicted of her second-strike will still receive double the typical sentence as mandated by the TSL ($f_j = F_{TSL}$), she now has the option to commit a non-strikeable offense in the future without getting sentenced to life in prison. Losing her second-strike is no longer as costly. Assuming that lenient counties had been effectively implementing Prop. 36 before it was formally passed, δ_{Lj} represents the risk of future punishment after the reform. Although this is a broad assumption, it is not unreasonable given that locking up repeat offenders for life due to a non-strikeable offense is one of the harshest aspects of the TSL. As a result, an offender's expected utility from committing a second strike after Prop. 36 will either increase if she is from a strict county or stay the same otherwise:

$$\Delta EU_{Sj,2nd} = p_j U_j\left(Y_{ji} - F_{TSLj} - \delta_{Lj}\right) - p_j U_j\left(Y_{ji} - F_{TSLj} - \delta_{Sj}\right) > 0 \tag{8}$$

$$\Delta EU_{Lj,2nd} = 0 \tag{9}$$

Equations (8) and (9) also hold when looking at the change in expected utility when an offender commits her first-strike offense. Losing her first-strike is less costly after the reform since the scope of offenses that result in harsher punishment in the future is more narrow. The only difference is that f_j will equal F_0 instead:

$$\Delta EU_{Sj,1st} = p_j U_j\left(Y_{ji} - F_{0j} - \delta_{Lj}\right) - p_j U_j\left(Y_{ji} - F_{0j} - \delta_{Sj}\right) > 0 \tag{10}$$

$$\Delta EU_{Lj,1st} = 0 \tag{11}$$

[53] Non-strikeable offenses by definition do not count as strikes if the offender has less than two prior strikes. Since they do not bring these offenders closer to harsher punishment and thus will not be affected by Prop. 36, we will not explicitly consider these cases. We will, however, consider the third-strike scenario where non-strikeable offenses do count as strikes.

Now, let us consider the case where an offender, who already has two previous strike convictions, commits a non-strikeable offense. Since it is assumed that Prop. 36 did not affect lenient counties, we will proceed to consider this case in relation to strict counties. Before Prop. 36 was passed, the offender in this example would have been sentenced to life in prison, as represented by F_{TSL}. Since this is her last strike, she no longer faces the threat of increased punishment in the future. Therefore, δ_{Sj} equals zero. After the reform of the TSL, the non-strikeable offense no longer counts as a third strike against the offender. According to Prop. 36, the offender now receives double the normal sentence for her current conviction instead of a life sentence. This will be represented by F_{DB}, where $F_{TSL} > F_{DB} > F_0$. In addition, the offender faces no risk of delayed punishment since her current conviction no longer constitutes a strike that moves her closer to more severe penalties. Therefore, δ_{Sj} also equals zero in the period after Prop. 36 was passed. The strict-county offender will thus have a higher expected utility after the reform:

$$\Delta EU_{Sj,3rdNS} = p_j U_j\left(Y_{ji} - F_{DBj}\right) - p_j U_j\left(Y_{ji} - F_{TSLj}\right) > 0 \tag{12}$$

$$\Delta EU_{Lj,3rdNS} = 0 \tag{13}$$

If the offender in the previous example had committed a strikeable offense instead, Prop. 36 would have had no impact on her expected utility, regardless of the county in which she resides. The offender would have still received a sentence of life in prison:

$$\Delta EU_{Sj,3rdS} = 0 \tag{14}$$

$$\Delta EU_{Lj,3rdS} = 0 \tag{15}$$

As seen from equations (12) and (14), relative to pre-reform, offenders facing their last strike after Prop. 36 (in a strict county) gain more utility from non-strikeable crimes than strikeable crimes. Assuming that these repeat offenders have a propensity towards crime, the reduced penalty on non-strikeable offenses after Prop. 36 could incentivize substitution away from strikeable crimes – those who would have committed a strikeable offense pre-reform may choose to switch to non-strikeable crimes post-reform. However, it is also possible that these offenders are hardened criminals who do not respond to incentives. In this case, they will not be deterred from committing strikeable offenses and there will be no substitution.

We have now considered all scenarios in which Prop. 36 could have affected the offender's decision to commit an offense. In each scenario, we show that, for strict counties after the reform, the offender's expected utility from both types of offenses either increases or stays the same. Offenders in lenient counties are assumed to be unaffected. Based on the standard crime model, the offender will participate in crime only if her expected utility from illegitimate activities exceeds what she gets from legitimate activities. Assuming that the latter does not change after the introduction of Prop. 36, the offender (in a strict county) will be more likely overall to participate in crime after the reform.

By limiting the harshest punishment to only strikeable offenses, Prop. 36 decreases the risk of more severe punishment in the future as well as the length of punishment itself for certain offenses. Thus, based on the presented model, Prop. 36 reduces the deterrent effect of the TSL, and we should expect to see the number of offenses, both strikeable and non-strikeable, increase in strict counties relative to lenient counties.[54]

5 Empirical Strategy

We use a difference-in-differences model in order to estimate the effect of Prop. 36 on various crime rates across California. In order to tease out this effect, we exploit the fact that counties varied in how strictly they implemented the TSL. We would expect the reform to affect strict counties more than lenient counties since the latter had been effectively implementing the reform before it was passed. Thus, strict counties represent the treatment group with Prop. 36 being the treatment, while lenient counties represent the control group, with “strict” and “lenient” as defined later in this section. By comparing differences in crime rates of strict and lenient counties before and after the reform’s introduction in 2012, we can tease out the effect of Prop. 36, assuming that in the absence of the reform, the change in crime rates in strict counties would have been the same as the change observed in lenient counties.

The baseline regression we run is as follows:

$$Y_{it} = a_i + \beta_1 Strict_{it} * Post_{it} + \beta_2 Z_{it} + \gamma_t + \theta_i + \epsilon_{it} \tag{16}$$

where Y_{it} is the crime rate of county i in year t, $Strict_{it}$ is a dummy variable equal to 1 if a county was strict in its implementation of the TSL, $Post_{it}$ is a dummy variable equal to 1 if the year is greater than 2012 (which indicates the period after Prop. 36 was passed), Z_{it} includes several economic control variables, and γ_t and θ_i are year and county fixed effects, respectively.

[54] We would only expect to see total strikeable offenses decrease in our model if substitution away from these offenses on the third strike exceeded the predicted increase from first- and second-strikers. This is unlikely since the number of people facing their first and second strike is much greater than the number of people facing their last strike.

The coefficient of interest that captures the causal effect of Prop. 36 is β_1, which represents the average change in crime rate in strict counties after the reform was introduced in 2012. County fixed effects control for variation in crime rate between strict and lenient counties due to their baseline differences. Year fixed effects account for other time varying factors that could affect the outcome. The diff-in-diff model controls for county and year fixed effects since they could bias our estimate of the average treatment effect on strict counties.

The effect that we identify with our diff-in-diff model is the pure deterrent effect of Prop. 36 because any incapacitation effect would be accounted for by county and year fixed effects. Suppose the TSL had affected crime rates partially through the incapacitation effect. In such a case, strict counties would experience a greater incapacitation effect than lenient counties (pre-reform) due to the substantial sentence enhancements provided for by the TSL. As a result, the change in crime rates over time would differ between counties and there would be differential trends. However, assuming that the incapacitation effect of the TSL varies over time but is fixed across strict counties and fixed across lenient counties, it would be accounted for by year fixed effects. Meanwhile, the difference between the incapacitation effect in strict and lenient counties, assuming that it is fixed over time, would be accounted for by county fixed effects. Thus, our diff-in-diff model controls for the incapacitation effect pre-reform.

Now suppose that Prop. 36 also had an incapacitation effect on crime. This will only affect our results if it causes strict counties to experience a different change in crime rates after the reform, relative to lenient counties. However, since strict and lenient counties impose the same sentencing policy after 2012, they will experience the same change in crime rates due to the incapacitation effect. As a result, our diff-in-diff model will not detect the incapacitation effect when comparing differences in crime rates of strict and lenient counties before and after the reform.

5.1 Identification Strategy

The critical assumption in diff-in-diff is that in the absence of treatment, the difference in crime between the treatment and control group is constant over time. Ideally, we would estimate the model by comparing counties in California with a control group consisting of counties from another three-strikes state that were not exposed to Prop. 36. However, as discussed in Section 2 of the paper, California's TSL was the harshest habitual offender law in the country. Due to the sheer number of offenses that trigger a strike sentence and the frequency with which California applied the TSL, it is not comparable to any other TSL state. Such a comparison would violate the parallel trends assumption of diff-in-diff.

Instead, we will define the treatment group as California counties that are strict in their application of the TSL while the control group as counties that are more lenient. Although lenient counties were also exposed to Prop. 36, they were effectively not treated because they had already been handling their TSL cases in a similar manner to what is mandated by Prop. 36. For example, some counties only applied a life sentence when cases were serious or violent, which is exactly what Prop. 36 mandated. Others handled them based on unrelated factors such as fiscal constraints, but we would still expect these counties to prioritize the most serious cases in practice when giving out life sentences.[55] As a result, we assume that lenient counties were generally not treated by the reform and would thus be a reasonable control group.

Assuming that crime patterns are uniform across counties, strictness is defined as the average percentage of second- and third-strikers among the prison population by county of commitment from 2004 to 2012. A higher percentage signals a stricter county that is sending more strikers to prison than other counties. Likewise, a lower percentage signals a more lenient county. Figure 3 shows the distribution of strictness among the 58 counties in California. The distribution is slightly skewed to the left, with a median of 20.17%. We split the counties based on whether they fall above or below the median strictness. We assume that this split is permanent, and that crime varies linearly with strictness.

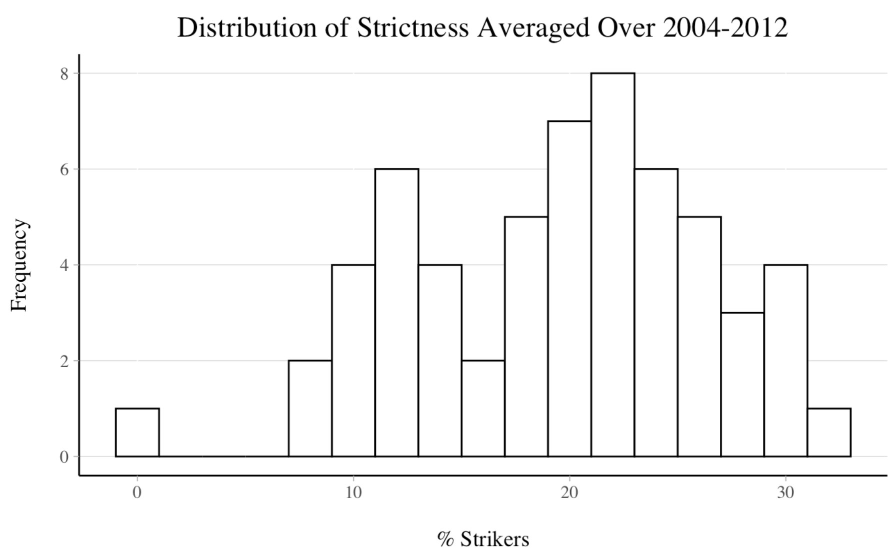


Figure 3: Distribution of strictness averaged over 2004-2012 for all 58 counties. Data is from California Department of Corrections and Rehabilitation (CDCR).

[55] Brown and Jolivette, ibid.

Table 2 below shows the summary statistics for the distribution of strictness. The first row includes data for all counties. The minimum is 0%, which represents the strictness of Alpine County. We consider it an outlier because Alpine County is California's least populous county, with a population of around 1,000 people and an average inmate population of 3 people. Strictness is thus irrelevant here due to the lack of crime being committed. As a result, we omit Alpine County from our analysis. As shown by the second row of Table 2, this leads to a median of 20.35%. Counties with strictness less than the median are considered lenient; otherwise, they are considered strict.

Table 2: Summary statistics of strictness averaged over 2004-2012

Variable	Median	Mean	St. Dev.	Min	Max	N
Strictness (all)	20.17	19.38	6.93	0.00	31.55	58
Strictness (excluding Alpine)	20.35	19.72	6.48	7.38	31.55	57

Source: California Department of Corrections and Rehabilitation (CDCR)

5.2 Control Variables

From the enhanced model of crime presented earlier, we see that Prop. 36 affects the individual's decision to offend through variables f_j and δ_j by changing the severity of punishment and the risk of delayed punishment, respectively. However, the individual's decision also depends on p_j (the probability of apprehension and conviction) as well as the legitimate income opportunities available. Thus, in order to isolate the effect of Prop. 36, the diff-in-diff model must control for these confounding factors.

Following Shepherd (2002), we split the probability of apprehension and conviction into two separate probabilities: the probability of arrest and the conditional probability of conviction given arrest. Since data on arrest is insufficient and data on convictions is not easily available, these two probabilities cannot be directly estimated. Instead, we will use the probability of clearance, *Prob(Cl)*, as a proxy for the probability of arrest, and the probability of imprisonment given clearance, *Prob(i|Cl)*, as a proxy for the probability of conviction given arrest. Both of these are calculated at the county level, the formulas of which are given in the Data section below.

We cannot control for these variables by directly adding them to our regression due to their simultaneous relationship with crime rates: more crime leads to more law enforcement measures but more law enforcement also results in fewer crimes.[56] However, if *Prob(Cl)* and *Prob(i|Cl)* are fixed over time, then the county fixed effect will absorb both their effect on crime and the simultaneity bias.

Figure 4 below show (a) the average *Prob(Cl)* and (b) the average *Prob(i|Cl)*, respectively, for each year from 2000-2017. We see that both variables remain relatively constant over time for all categories of crime. The drop in average *Prob(i|Cl)* in 2012 is due to the Public Safety Realignment Act that was introduced in October 2011. Realignment reduced the number of inmates in state prison by mandating that all non-violent, non-serious, and non-sex offenders ("triple-non") serve their sentences in county jails rather than state prisons.[57, 58] This reduction shows up in Figure 4(b) because *Prob(i|Cl)* is calculated based on the number of offenders in state prisons; it does not include data from county jails. But since Realignment applied to all counties in California, this drop is absorbed by the year fixed effect. Thus, Figure 4(a) and Figure 4(b) suggest that it is reasonable to assume that *Prob(Cl)* and *Prob(i|Cl)* are fixed over time and absorbed by the county fixed effect for all categories of crime.

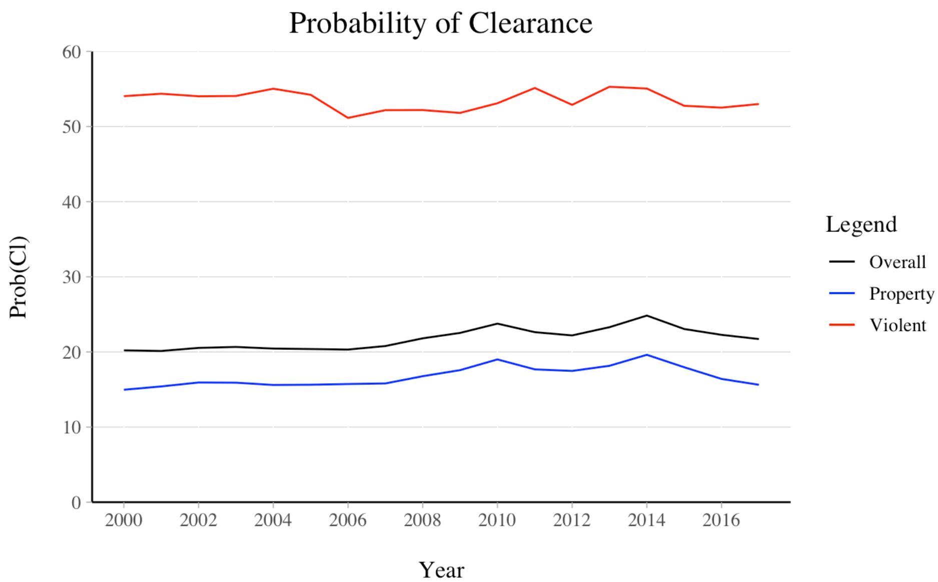
We also assume that neither *Prob(Cl)* nor *Prob(i|Cl)* is correlated with the introduction of Prop. 36. This is reasonable to assume since it is unlikely that changes in law enforcement at the county level would affect California's decision to reform the TSL. Under this assumption, the effect of Prop. 36 on crime will not be confounded by *Prob(Cl)* or *Prob(i|Cl)*.

Finally, the legitimate income opportunities available to an offender will be controlled for through several economic control variables. These variables include per capita personal income, per capita unemployment insurance payments, and per capita income maintenance payments.

[56] Isaac Ehrlich, "Crime, Punishment, and the Market for Offenses," *Journal of Economic Perspectives* 10, no. 1 (February 1, 1996): 43–67, https://doi.org/10.1257/jep.10.1.43.

[57] Dean Misczynski, "Corrections Realignment: One Year Later," *Public Policy Institute of California*, August 2012, 38.

[58] Realignment did not transfer any current inmates already in prison. Instead, all provisions were applied prospectively. However, the "triple-non" requirement applies to both current and prior convictions.



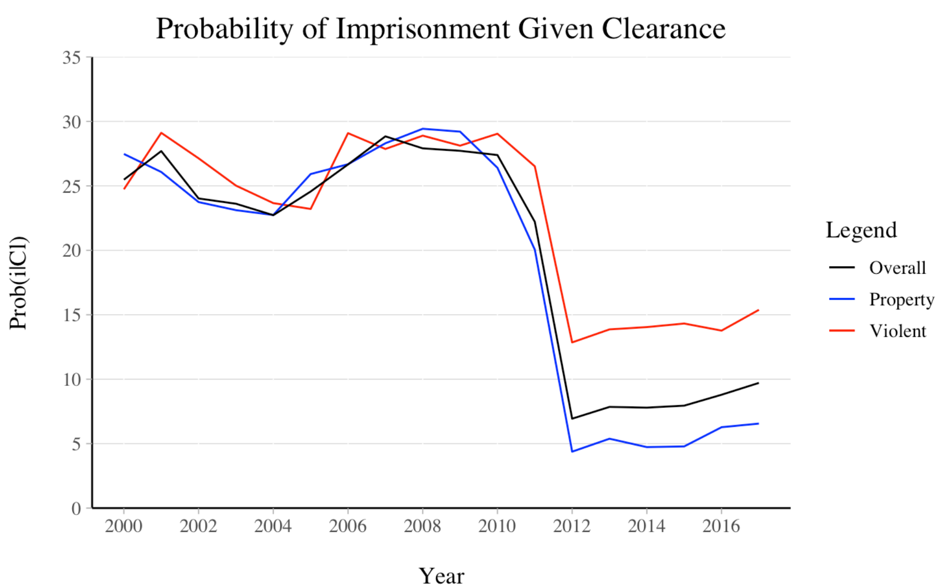


Figure 4: These plots show (a) the average probability of clearance and (b) the average probability of imprisonment given clearance from 2000-2017. For each year, the average is taken across all counties except Alpine County. Data is from the Bureau of Justice Statistics' National Corrections Reporting Program (NCRP).

6 Data

6.1 Data Source

We use county-level data on 57 California counties (excluding Alpine County, see discussion above) between 2000-2017. We consider 2013 as the first year that Prop. 36 was implemented, which allows thirteen years of data in the period before Prop. 36 and five years in the period after.[59] County dummies capture differences across counties that are fixed over time, such as law enforcement patterns and differences in strictness. Year dummies account for trends in crime that vary over time.

Crime data is provided by the Federal Bureau of Investigation's Uniform Crime Reporting Program (UCR).[60] The UCR Program was created in 1930 to consolidate crime statistics from law enforcement agencies across the country. In order to ensure that the data is uniformly reported, the UCR Program provides agencies with standardized definitions of offenses. It collects data on eight "index crimes" and splits them into two categories: violent crime and property crime. Violent crime is defined as crime involving force or threat of force and includes homicide, rape, robbery, and aggravated assault. On the other hand, property crime includes burglary, motor vehicle theft, larceny-theft, and arson.[61]

In multiple-offense incidents, the UCR Program standardizes the reporting procedure with the Hierarchy Rule, which states that law enforcement agencies must only report the offense that is the most serious as defined by the hierarchy list. Violent crimes are listed as more serious than property crimes.[62] The Hierarchy Rule does not apply to arson since arson is often committed in conjunction with another index crime. In such cases, both arson and the other crime are counted.

We consider three different crime rates as our outcome variables: violent crime, property crime, and overall crime. Violent crime in our regression includes rape, robbery, and aggravated assault. These crimes are all considered strikeable offenses under the TSL. We omit homicide since there are far fewer homicides reported than other violent offenses. Property crime in our regression includes burglary, motor vehicle theft, and larceny-theft. Most of these crimes are considered

---

[59] Although Prop. 36 was enacted immediately after it was passed in November 2012, county-level data is only available by year, so we will consider 2013 as the beginning of the post-reform period for analysis purposes.

[60] U.S. Department of Justice. Federal Bureau of Investigation. Uniform Crime Reporting Program. Retrieved from https://crime-data-explorer.fr.cloud.gov/explorer/state/california/crime

[61] U.S. Department of Justice, Federal Bureau of Investigation. Uniform Crime Reporting Handbook : UCR. [Washington, D.C.] :U.S. Dept. of Justice, Federal Bureau of Investigation, 2004.

[62] The hierarchy list is as follows in descending order of seriousness: homicide, rape, robbery, aggravated assault, burglary, larceny-theft, and motor vehicle theft.

non-strikeable offenses under the TSL.[63] We omit arson for consistency purposes since it does not follow the Hierarchy Rule. Overall crime is defined as the sum of violent crime and property crime (as defined in our regression). We choose to focus on these broader categories of crime instead of index crimes individually so as to avoid problems related to multiple hypothesis testing.

The crime rate is defined as crime per 100,000 people. It is calculated for each county as the total number of offenses reported divided by county population in units of 100,000 people. Data on the number of offenses reported is pulled from the Criminal Justice Statistics Center (CJSC) in the California Department of Justice, which collects data according to the standards specified by the UCR Program. County population data comes from the California Association of Counties (CSAC).[64, 65]

We calculate a county's strictness by first estimating, for each county, the percentage of second- and third-strike offenders among its prison population each year (the number of second- and third-strike offenders in prison divided by the total prison population).[66] Data on the number of strike inmates as well as total prison population are obtained from the California Department of Corrections and Rehabilitation (CDCR).[67] To find each county's strictness level, we then average its strictness over the period 2004-2012.[68] If this average falls below the median, we set *Strict* equal to 1. Otherwise, we set the variable equal to 0.

The probability of clearance is estimated with the clearance rate. As defined by the UCR Program, "clearance" occurs when at least one person involved in the

---

[63] Not all property crimes are non-strikeable offenses. For example, residential burglary counts as a strikeable offense while commercial burglary does not.

[64] U.S. Department of Justice. Office of the Attorney General. Criminal Justice Statistics Center. Crimes and Clearances. Retrieved from https://openjustice.doj.ca.gov/data

[65] California State Association of Counties. Population Estimates for Counties and Cities – 1970 to 2018. Retrieved from https://www.counties.org/post/datapile

[66] A county's prison population is defined by county of commitment. For example, those serving time for crimes committed in Alameda County are considered as Alameda's prison population.

[67] California Department of Corrections & Rehabilitation. Second and Third Striker Felons in the Adult Institution Population, Prison Census Data. Retrieved from https://www.cdcr.ca.gov/research/research-requests/

[68] The average starts in 2004 because that was the year that the CDCR changed the way it counted second- and third-strike offenders: it began excluding those who are returned to custody for a parole violation. The average ends in 2012 since that is the last year before the introduction of Prop. 36.

offense is (1) arrested, (2) charged, and (3) turned over to the court for prosecution.[69] For each category of crime, the clearance rate is calculated as the number of clearances divided by the number of offenses reported in each county.[70] The data comes from the CJSC.

The probability of imprisonment given clearance is calculated for each category of crime by dividing the number of offenders sentenced to prison by the number of clearances. The data on the number of offenders entering prison is obtained from the Bureau of Justice Statistics' National Corrections Reporting Program (NCRP).[71] Since 1983, the NCRP has compiled offender-level data on prisoner admissions and releases across the country. Individual inmate records are collected annually at the county level and includes information such as the offense(s) that the prisoner is serving time for, the county in which the offense was committed, and the date of admission to state prison.

In cases where the offender is serving time for multiple crimes, the NCRP reports up to three offenses that are associated with that particular sentence. This is different from how clearances are reported by the UCR Program, which uses the Hierarchy Rule when dealing with cases involving multiple crimes. In order to compare apples to apples when dividing the number of sentenced offenders by the number of clearances for a particular crime, we apply the Hierarchy Rule to prison admission data; we only count the most serious crime that the NCRP lists.[72]

The economic variables that we control for include per capita personal income, per capita unemployment insurance payments, and per capita income maintenance payments. The data for these variables comes from the Regional Economic Accounts of the Bureau of Economic Analysis. Personal income represents the wealth provided by legitimate activities as well as the wealth available for potential

---

[69] Since clearance requires that the arrested offender is also charged and turned over to the court for prosecution, the probability of clearance as a proxy for the probability of arrest may underestimate the latter.

[70] In a few instances, the clearance rate came out to be greater than 1, meaning that there were more clearances than the number of crimes reported. Since only reported offenses are cleared, this value does not make sense. Thus, in such instances, we set the value of PC_{it} to NA and treat it as missing.

[71] United States. Bureau of Justice Statistics. National Corrections Reporting Program, [United States], 2000-2017. Inter-university Consortium for Political and Social Research [distributor], 2020-11-19. https://doi.org/10.3886/ICPSR37608.v1

[72] In cases where an offender is readmitted to prison, the NCRP reports both the crime(s) that the offender had previously committed and the new crime(s) that she is currently serving time for. As such, it is impossible to know exactly which crime(s) a repeat offender is getting sentenced for. For consistency's sake, we only consider the crimes for which the repeat offender has not previously been sentenced. In cases where all of the listed crimes had been committed before, we consider all of them when applying the Hierarchy Rule.

offenders to steal. Unemployment insurance payments measures job availability and the ease with which potential criminals can enter the job market. Income maintenance payments are welfare benefits available to those who are poor and unemployed, and it controls for legitimate sources of income outside of the job market.

6.2 Summary Statistics

Table 3 shows the summary statistics for data averaged over the period 2000-2017. The data is split between strict and lenient counties in order to capture how they differ in crime rates, probabilities of clearance, probabilities of imprisonment given clearance, and various economic factors. As mentioned earlier, we omit Alpine County as an outlier. Among the remaining 57 California counties, 28 counties are considered strict while 29 counties are considered lenient.

On average, strict and lenient counties do not significantly differ in either the violent crime rate or the property crime rate. However, offenders in strict counties are less likely to get cleared for violent crime. Furthermore, they are also less likely to go to prison if they are cleared for property crime. With regards to overall crime, offenders in strict and lenient counties do not face significant differences in terms of either probability of clearance or probability of imprisonment if cleared.[73]

Economic factors differ significantly between strict counties and lenient counties. Strict counties have, on average, higher income, lower unemployment, and less people on welfare. These differences indicate that strict counties are better off economically than lenient counties.

[73] *Prob(Cl)* and *Prob(i|Cl)* for overall crime are both statistically insignificant at the 5% significance level.

Table 3: Summary statistics

Variable	Strict Mean	Lenient Mean	Difference	t-statistic
Violent Crime				
Crime rate	436.43	421.56	14.87	1.36
Prob(Cl)×100	51.32	55.57	-4.26	-4.56
Prob(i|Cl)×100	23.32	22.34	0.98	1.16
Property Crime				
Crime rate	2,761.49	2,719.91	41.58	0.62
Prob(Cl)×100	16.38	17.08	-0.70	-1.24
Prob(i|Cl)×100	18.77	20.92	-2.15	-2.21
Overall Crime				
Crime rate	3,197.92	3,141.47	56.45	0.76
Prob(Cl)×100	20.98	22.51	-1.53	-2.51
Prob(i|Cl)×100	19.82	21.44	-1.62	-1.75
Economic				
Personal income	41,165.39	38,514.48	2,650.91	2.78
Unemployment payments	242.54	274.39	-31.84	-3.03
Welfare payments	697.19	790.54	-93.35	-4.92

Note: Mean is calculated as data averaged over the period 2000-2017. Lenient counties (N=29) and strict counties (N=28) have strictness below and above the median of 20.35%, respectively. Difference is calculated as Strict Mean minus Lenient Mean. Crime rate is given as number of offenses per 100k people. Prob(Cl) and Prob(i|Cl) are given as percentages. Economic variables are given in per capita units. t-statistics are calculated using robust standard errors that are clustered at the county level.

7 Results

7.1 Comparison of Crime Rates

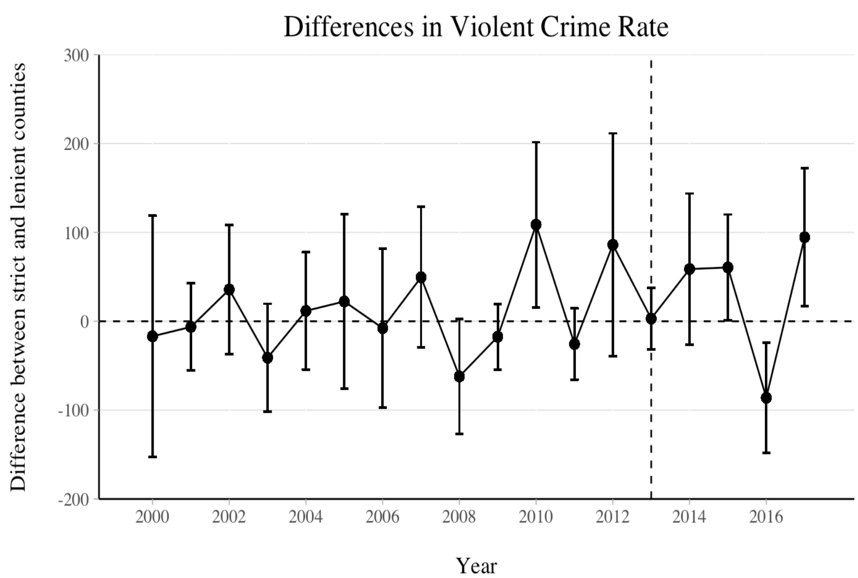
The diff-in-diff model estimates the effect of Prop. 36 on crime based on the assumption that the counterfactual growth in crime for strict counties is the same as the growth observed in lenient counties. Although we cannot prove this assumption, we can determine its plausibility by looking at crime trends prior to the reform. If both groups experience the same growth pre-reform, then it is more likely for there to be parallel trends in the absence of treatment.

Figure 5 plots the differences in average crime rates between strict counties and lenient counties from 2000-2017. The plots correspond to (a) violent crime, (b) property crime, and (c) overall crime, respectively. The vertical dashed line marks the year 2013 in which Prop. 36 was introduced. If strict and lenient counties have common trends pre-reform, we would expect to see equal differences in means prior to 2013.

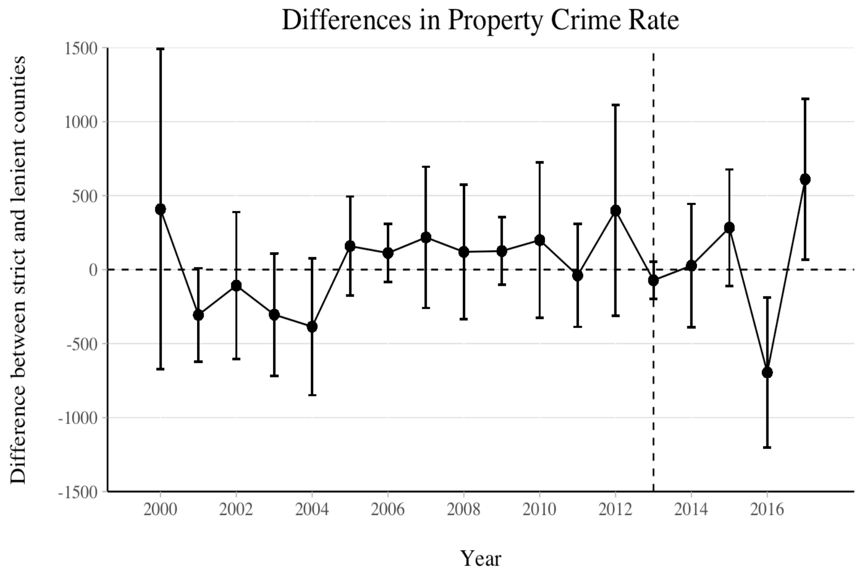
Figure 5(a) shows that in 2010, strict counties had significantly higher rates of violent crime than lenient counties, where significance is defined at the 5% level. Besides that, there are no other significant differences before the reform. This suggests that for violent crime, pre-trends might differ between strict and lenient counties.

Property crime and overall crime show very similar plots, which is likely due to the fact that the majority of crimes committed are property crimes.[74] In both of their plots, there appears to be no significant differences in crime rates prior to 2013. Significance is defined at the 5% level. This suggests that counties have the same pre-trend for property crime and overall crime, which means that the parallel trends assumption is likely to hold.

If Prop. 36 had an effect on crime, we would expect the differences in average crime rates to reflect that by diverging from their constant pre-reform levels – such a divergence, whether it is positive or negative, would suggest a change in crime rates that is potentially due to the reform. As we see in Figure 5, in 2016, strict counties had significantly lower crime rates than lenient counties for all categories of crime. However, they then experienced significantly more crime than lenient counties in the following year. Because these differences in means do not move in any particular direction for any category of crime, there are no discernible effects of Prop. 36 on crime.



[74] Overall crime is defined as violent crime plus property crime. In 2012, 86% of overall crime consisted of property crime.



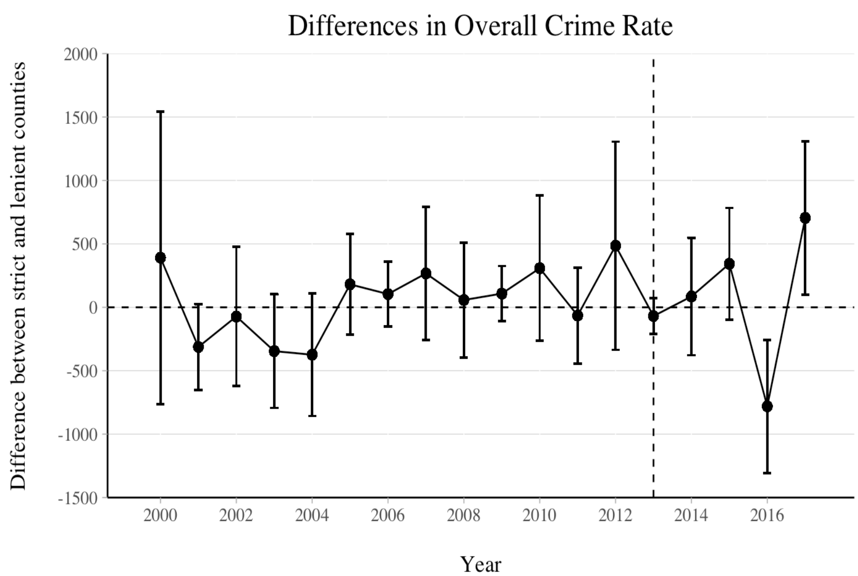


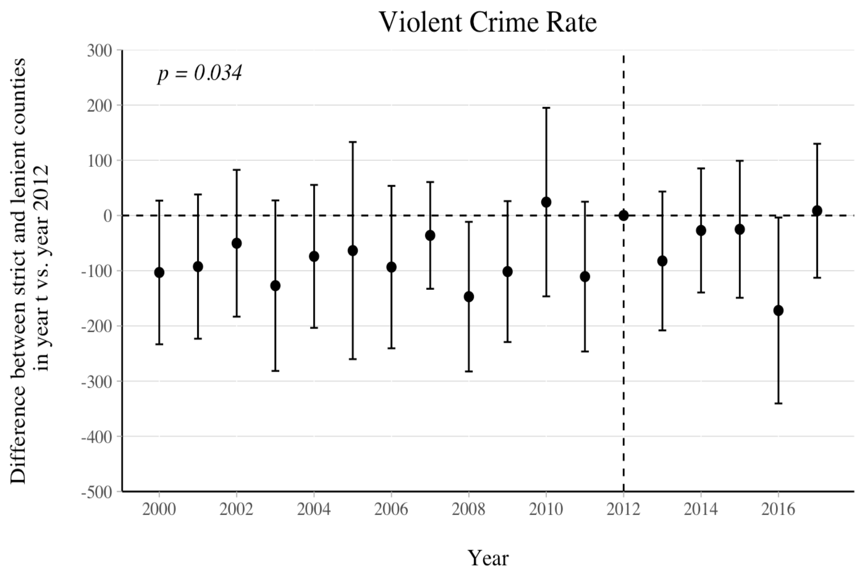
Figure 5: Differences in means of (a) the violent crime rate, (b) the property crime rate, and (c) the overall crime rate between strict and lenient counties. The crime rate is given as crime per 100,000 people. Alpine County is excluded. The vertical dashed line marks the year 2013 in which Prop. 36 was introduced. The vertical bars correspond to 95% confidence intervals, which are calculated with robust standard errors clustered at the county-level.

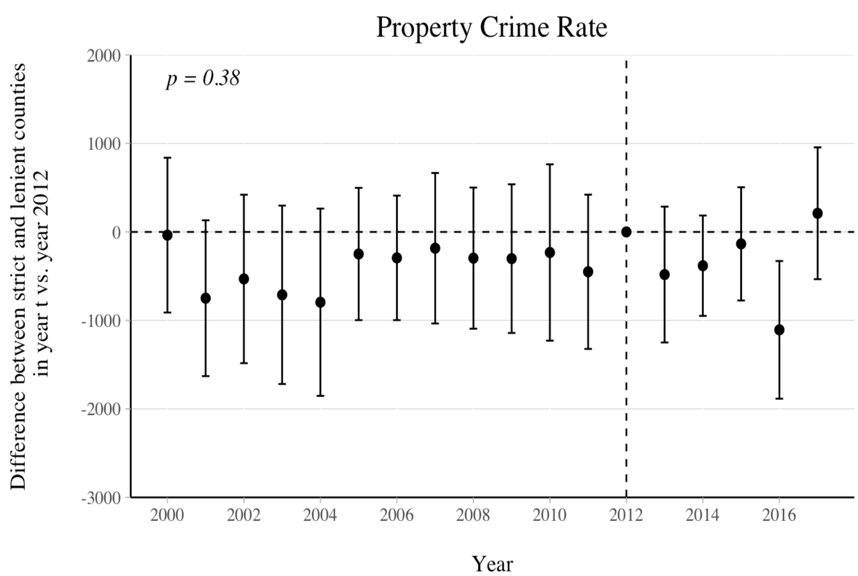
7.2 Testing for Differential Trends

We now present a more rigorous test for determining whether the parallel trends assumption holds. We run a separate regression for each category of crime to test for differential trends. The regression used is similar to equation (16) in Section 5, but instead of *Strict***Post*, we interact *Strict* with a dummy variable for every year besides 2012 (the last year before the reform). The coefficients on the interaction terms provide insight into whether crime trends differ between counties – they compare the difference in crime rates between strict and lenient counties in the corresponding year against the same difference in 2012. The pre-reform interaction effects tell us whether any pre-trends exist. In order to address the issue of multiple hypothesis testing, we jointly test that none of the coefficients corresponding to the pre-period are significant. The results are reported in Appendix Table A1.

Figure 6 plots the coefficients of interest for (a) violent crime, (b) property crime, and (c) overall crime, respectively. If there are no differential trends before the reform, we would expect the coefficients before 2012 (marked by the vertical dashed line) to not be significantly different. This would indicate that the differences in crime rates are the same for each year prior to the reform. The joint test tells us whether these differences are equal to zero at the same time. A significant result indicates the existence of differential trends. The p-value for each joint test is shown on the upper-left corner of the corresponding plot.

Figure 6(a) shows that for violent crime, the only interaction effect that is significant at the 5% level is the one corresponding to the year 2008. This result suggests that strict and lenient counties have different pre-trends for violent crime. On the other hand, property and overall crime appear to have the same pre-trends – none of their interaction effects before 2012 are significantly different from zero. These results are consistent with the earlier graphical comparison of crime rates. From the joint tests, we also see that the only significant result at the 5% level is for violent crime ($p = 0.034$). Thus, the joint tests confirm that there are differential trends for violent crime but not for either property crime ($p = 0.38$) or overall crime ($p = 0.42$).


Violent Crime Rate
p = 0.034
Difference between strict and lenient counties in year t vs. year 2012
300
200
100
0
-100
-200
-300
-400
-500
2000
2002
2004
2006
2008
2010
2012
2014
2016
Year


Property Crime Rate
p = 0.38
Difference between strict and lenient counties in year t vs. year 2012
2000
1000
0
-1000
-2000
-3000
2000
2002
2004
2006
2008
2010
2012
2014
2016
Year

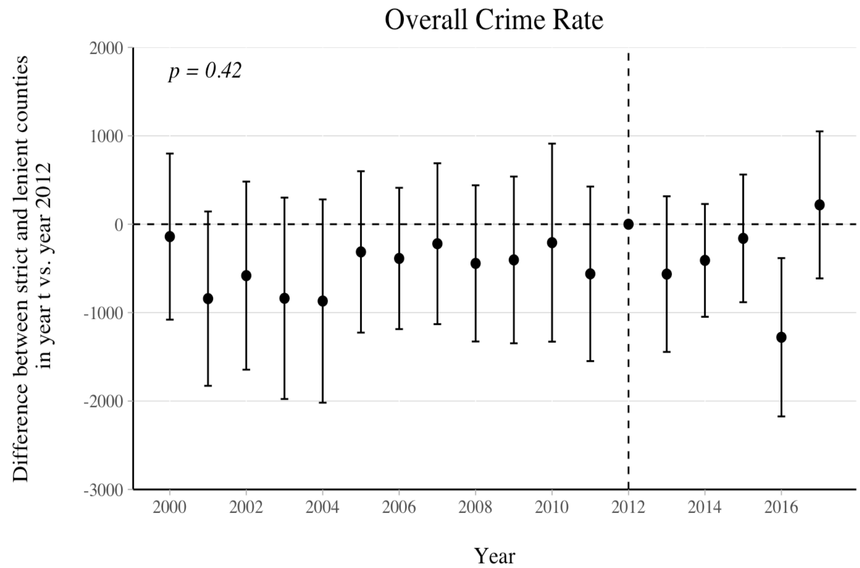


Figure 6: Differential trends test for (a) the violent crime rate, (b) the property crime rate, and (c) the overall crime rate. Each point represents the difference in crime rate between strict and lenient counties in year t vs. year 2012. The crime rate is given as crime per 100,000 people. Alpine County is excluded. The vertical dashed line marks the year 2012, the last year before Prop. 36 was introduced. The vertical bars correspond to 95% confidence intervals, which are calculated with robust standard errors clustered at the county-level. The p-value refers to the joint test that none of the pre-period (2000-2011) interaction effects are significant.

7.3 Diff-in-Diff Results

Table 4 reports the results of the diff-in-diff regression specified in SectionX. The regression is run separately for each outcome variable: violent crime, property crime, overall crime. The coefficient of interest is on the interaction term *Strict*Post*, which estimates the average treatment effect of Prop. 36 on strict counties.

Table 4: Baseline diff-in-diff estimates

	Violent	Property	Overall
	(1)	(2)	(3)
Strict×Post	−23.34	−7.49	7.91
	(−69.72, 23.03)	(−245.63, 230.65)	(−238.95, 254.77)
Strict×Trend	Y	N	N
Number Years	18	18	18
Number Counties	57	57	57
Observations	1,026	1,026	1,026

Note: Each column estimates equation (16) for the corresponding category of crime. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Counties are grouped based on strictness averaged over 2004-2012. The row *Strict×Trend* represents whether the corresponding regression controls for differential trends. Coefficients for economic controls, differential trend controls, and fixed effects are not shown. The sample used to estimate each regression includes all counties except Alpine County. The 95% confidence interval, calculated with robust standard errors clustered at the county-level, is given in parentheses.

We control for differential trends in violent crime by adding to its regression an interaction term between a dummy variable for strictness and a linear trend variable, represented in Table 4 as *Strict*Trend*.[75] We assume crime grows over time in a linear fashion. This approach is consistent with what others in the TSL literature have previously used to control for underlying crime trends.[76] After adding this new interaction term, we see in column (1) of Table 4 that the coefficient on *Strict*Post*

[75] The linear trend variable, *Trend*, is defined as (Year – 2013), where 2013 is the first year of treatment.

[76] Worrall, ibid; Elsa Y. Chen, "Impacts of 'Three Strikes and You're Out' on Crime Trends in California and Throughout the United States," *Journal of Contemporary Criminal Justice* 24, no. 4 (November 1, 2008): 345–70, https://doi.org/10.1177/1043986208319456.

is not significant. Based on the 95% confidence interval, Prop. 36 could have decreased the violent crime rate in strict counties by as much as 11.1% or increased it by up to 3.7%, relative to the average violent crime rate in 2012.[77]

Column (2) corresponds to the regression for property crime. The coefficient of interest is also not significant. The 95% confidence interval shows that relative to 2012, Prop. 36 could have decreased the property crime rate in strict counties by as much as 6.3% or increased it by up to 5.9%.[78]

In column (3), we see that the results for overall crime is also not significant, which is not surprising given that overall crime is just the combination of violent crime and property crime. The 95% confidence interval suggests that Prop. 36 could have decreased the overall crime rate in strict counties by as much as 5.3% relative to 2012. On the flip side, it also could have increased the overall rate by up to 5.6%.[79]

Therefore, the diff-in-diff results suggest that the TSL did not decrease violent crime like it had intended, nor did it decrease property crime. If the TSL were effective in decreasing either of those, we would expect to see a significant increase in their respective crime rates after the reform. Instead, we see that Prop. 36 had no significant impact on either crime rates.

The current regression assumes that crime varies linearly with strictness. If this assumption is correct, then our estimates will not depend on the counties we compare. However, if this assumption is false and crime actually varies nonlinearly with strictness, then differences in crime will depend on which counties we compare. In this case, our estimates will be incorrect – splitting the counties into two groups and comparing their average crime rates will not give us their true difference in crime. Thus, our current regression does not allow for flexibility in the relationship between crime and strictness. We now turn to an exploration of nonlinearity.

7.4 Exploration of Nonlinearity

We will explore the possibility that crime varies nonlinearly with strictness by dividing the 57 counties into four strictness buckets: super-lenient, medium-lenient, medium-strict, and super-strict. This allows us to get a better understanding of how Prop. 36 affects crime conditional on county strictness. The regression we now run is as follows:

$$Y_{it} = a_i + \beta_1 MediumLenient_{it} * Post_{it} + \beta_2 MediumStrict_{it} * Post_{it} + \beta_3 SuperStrict_{it} * Post_{it} + \beta_4 Z_{it} + \gamma_t + \theta_i + \epsilon_{it} \tag{17}$$

---

77 The average violent crime rate in 2012 was 630 crimes per 100,000 people.

78 The average property crime rate in 2012 was 3,894 crimes per 100,000 people.

79 The average overall crime rate in 2012 was 4,524 crimes per 100,000 people.

where $MediumLenient_{it}$, $MediumStrict_{it}$, and $SuperStrict_{it}$ are dummies equal to 1 if a county was medium-lenient, medium-strict, or super-strict in its implementation of the TSL, respectively. All other variables are as defined in the baseline regression in Section 5 above. Like before, the outcomes of interest are violent crime, property crime, and overall crime.

We define strictness buckets based on the distribution of strictness presented earlier. Table 5 below shows the relevant summary statistics describing that distribution. Counties are categorized relative to where their strictness falls in comparison to the first quartile, the median, and the third quartile. We continue to omit Alpine County as an outlier.

Table 5: Summary statistics of strictness averaged over 2004-2012

Statistic	Min	p25	Median	p75	Max	N
Strictness (all)	0.00	13.71	20.17	24.11	31.60	58
Strictness (excluding Alpine)	7.38	13.98	20.35	24.12	31.60	57

Source: California Department of Corrections and Rehabilitation (CDCR)

Using the values in the second row of Table 5, the strictness buckets are defined as follows: super-lenient counties are those whose strictness fall below the first quartile, medium-lenient counties are those between the first quartile and the median, medium-strict counties are those between the median and the third quartile, and super-strict counties are those above the third quartile. With this definition, we end up with 15 super-lenient counties and 14 counties in each of the remaining categories.

The coefficients of interest are the interaction effects, β_1, β_2, and β_3. They represent the average effect of Prop. 36 on crime rates in medium-lenient, medium-strict, and super-strict counties relative to super-lenient counties, respectively. Here, super-lenient counties are considered the control group that is assumed to not be affected by Prop. 36. In order to address the issue of multiple hypothesis testing, we jointly test that none of these interaction effects are significant.

The interaction effects estimate the causal effect of Prop. 36 if the parallel trends assumption holds for counties in each strictness bucket relative to super-lenient counties. We test for differential trends using the same method described earlier in Section 7.2. For each category of crime, we run three separate regressions to test whether pre-trends differ between super-lenient counties and counties in each of the other three strictness buckets. The results are reported in Appendix Table A2-A4.

Table 6 below summarizes the results of the joint tests. Each column corresponds to a different outcome variable. We can see that for all categories of crime, medium-lenient, medium-strict, and super-strict counties all have significantly different pre-trends relative to super-lenient counties. This suggests that the parallel trends assumption does not hold and that our results should be interpreted with caution.

Table 6: Joint test p-value

Strictness	Violent	Property	Overall
MediumLenient	< 0.001	< 0.001	0.002
MediumStrict	< 0.001	0.004	0.005
SuperStrict	0.039	0.034	0.020

Note: Each row corresponds to separate tests for differential trends between the counties represented by that row and super-lenient counties. All p-values refer to the corresponding joint test that none of the pre-period interaction effects are significant.

In order to control for differential trends, we interact a dummy variable for each strictness bucket with a linear time trend variable: *MediumLenient*Trend*, *MediumStrict*Trend*, *SuperStrict*Trend*.[80] Table 7 shows the results for the regression specified earlier in this section after controlling for differential trends. Each column corresponds to a separate regression: violent crime, property crime, and overall crime, respectively.

From column (1) in Table 7, we see that the coefficient on *SuperStrict*Post* is significant at the 5% significance level. This suggests that on average, Prop. 36 decreased the violent crime rate in super-strict counties by 11.5%, relative to the average violent crime rate in 2012. Prop. 36 appears to have no impact on the violent crime rate in other counties. It is interesting that violent crime only decreased significantly in super-strict counties after the reform. If these counties had been less strict, like medium-strict or medium-lenient counties, then we would expect them to have seen no significant changes in their crime rates following the reform. Thus, in terms of violent crime, the effect of Prop. 36 appears to depend on how strictly a county applied the TSL.

However, despite the significant coefficient on *SuperStrict*Post*, the joint test shows that Prop. 36 had no significant effect on the violent crime rate in any of the counties ($p = 0.10$). This result is also consistent with the insignificant result we saw in the baseline regression. Thus, we should view the significant result we see here for super-strict counties with some caution.

[80] The linear trend variable, *Trend*, is defined as (Year – 2013), where 2013 is the first year of treatment.

Column (2) and column (3) correspond to the regressions for property crime and overall crime, respectively. Neither the interaction effects nor the joint tests are significant (p = 0.64 for property crime and p = 0.56 for overall crime), which is consistent with our results from the baseline regression. However, the 95% confidence intervals here are much wider than what we saw earlier. This indicates that we do not have enough statistical power here to conclude whether property crime or overall crime were affected by Prop. 36. We should thus view our results here with caution.

Table 7: Exploration of nonlinearity

	Violent	Property	Overall
	(1)	(2)	(3)
MediumLenient×Post	−44.83	11.48	−33.35
	(−115.69, 26.03)	(−743.61, 766.56)	(−817.66, 750.96)
MediumStrict×Post	−22.73	−180.99	−203.73
	(−71.60, 26.13)	(−968.64, 606.65)	(−1,011.73, 604.27)
SuperStrict×Post	−72.17**	−380.80	−452.97
	(−133.51, −10.82)	(−1,117.29, 355.68)	(−1,201.02, 295.08)
MediumLenient×Trend	Y	Y	Y
MediumStrict×Trend	Y	Y	Y
SuperStrict×Trend	Y	Y	Y
***p*-value**	**0.10**	**0.64**	**0.56**
Number Years	18	18	18
Number Counties	57	57	57
Observations	1,026	1,026	1,026

Note: Each column estimates equation (17) for the corresponding category of crime. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Counties are grouped based on strictness averaged over 2004-2012. The rows *MediumLenient×Trend*, *MediumStrict×Trend*, and *SuperStrict×Trend* represent whether the regression controls for the corresponding differential trends. The row *p-value* refers to the joint test that none of the coefficients of interest are significant. Coefficients for economic controls, differential trend controls, and fixed effects are not shown. The sample used to estimate each regression includes all counties except Alpine County. The 95% confidence interval, calculated with robust standard errors clustered at the county-level, is given in parentheses.
*p < 0.1, **p < 0.05, ***p < 0.01

It is interesting to note, however, that the 95% confidence intervals for property crime consistently shift down with increasing strictness.[81] For example, relative to 2012, property crime in medium-lenient counties can increase by up to 19.7% after the reform, while in medium-strict counties, it can increase by up to only 15.6%. For super-strict counties, the upper bound is even lower with an increase of up to 9.1%. This suggests that in the worst-case scenario, the detrimental effect of Prop. 36 on the property crime rate goes down with a county's strictness.

7.5 Robustness Checks

From 1980 to 2006, California's state prison population dramatically increased, leading to severe overcrowding and unsafe prison conditions.[82] "Tough on crime" policies dating back to the 1970s and longer mandatory sentencing under the TSL contributed to this explosion. As discussed in Section 2, the federal court ordered the state of California in 2009 to reduce its prison population to 137.5% of design capacity.[83] At the time, this order translated into a reduction of about 44,000 prisoners. Since then, California has pursued many prison reforms in order to bring its prison population in line with the federal mandate.

The push for criminal justice reforms and the public's generally negative sentiment towards mass incarceration could have encouraged counties to apply the TSL less frequently. This would cause strictness to be endogenous to the onset of the reform since it is estimated with data from 2004-2012, which includes this period of increased criminal justice reform. If this is the case, then it would confound our earlier results.

We perform a robustness check by estimating strictness with data from 2004-2008 since 2008 is the last year before California was ordered to reduce its prison population. That is, for each county, we average the percentage of second- and third-strikers among its prison population over the period 2004-2008. Figure 7 below shows the distribution of strictness among the 58 counties based on this new estimation. The distribution is relatively normal, with a median of 17.86%.

---

[81] Overall crime shows a similar pattern with its confidence intervals, but this is most likely due to the fact that property crime account for the vast majority of crime overall.

[82] Magnus Lofstrom, Mia Bird, and Brandon Martin, "California's Historic Corrections Reforms," *Public Policy Institute of California* (blog), September 2016, https://www.ppic.org/publication/californias-historic-corrections-reforms/.

[83] Paul Golaszewski, "A Status Report: Reducing Prison Overcrowding in California," Legislative Analyst's Office, August 5, 2011, https://lao.ca.gov/reports/2011/crim/overcrowding_080511.aspx.

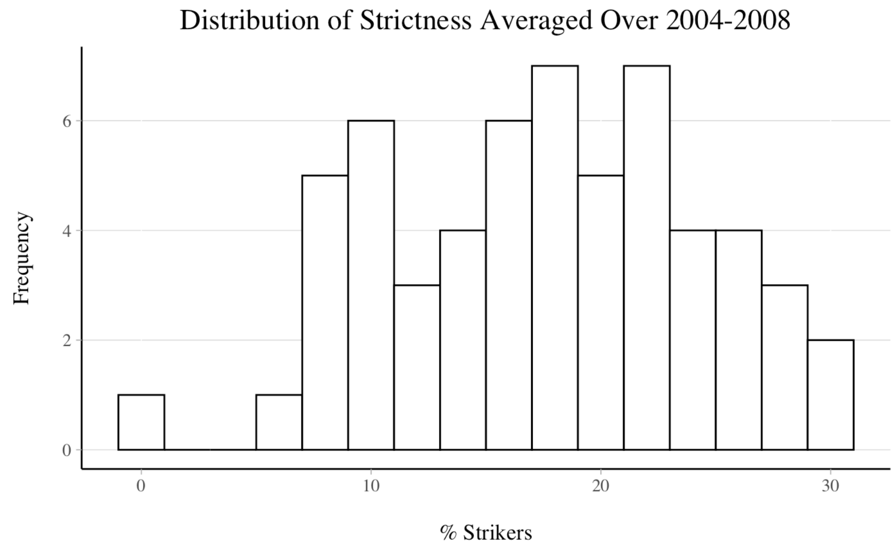


Figure 7: Distribution of strictness averaged over 2004-2008 for all 58 counties. Data is from California Department of Corrections and Rehabilitation (CDCR).

Table 8 shows the summary statistics for the above distribution. The first row includes data for all counties. Again, Alpine County is considered an outlier, and like before, we omit it from our analysis. As shown by the second row of Table 8, this leads to a median of 17.9%. Counties with strictness less than 17.9% are considered lenient; otherwise, they are considered strict. This gives us 28 lenient counties and 29 strict counties. Furthermore, under the robust definition of strictness, Madera and Yolo County are now considered lenient; Contra Costa, Nevada, and Santa Barbara County are now considered strict.

Table 8: Summary statistics of strictness averaged over 2004-2008

Statistic	Median	Mean	St. Dev.	Min	Max	N
Strictness (all)	17.86	17.57	6.81	0.00	30.10	58
Strictness (excluding Alpine)	17.90	17.88	6.45	6.39	30.10	57

Source: California Department of Corrections and Rehabilitation (CDCR)

Compared to the mean of the baseline strictness distribution (19.38%), this distribution based on the robust definition of strictness has a lower mean (17.57%). One possible reason for this is that counties applied the TSL more frequently when sentencing offenders in the four years leading up to Prop. 36. However, it is also expected that the mean goes up when we average across more years (2004-2012 instead of 2004-2008) because offenders sentenced under the TSL receive a substantially longer prison term such that there are more strikers entering prison than leaving at any given time. This means the total number of strikers in prison increases over time, which increases the percentage of strikers that make up the prison population, holding all else constant.

Table 9 reports the results of the robustness checks, where we estimate the baseline diff-in-diff regression with the new set of treatment and control groups here. We do not control for differential trends because strict and lenient counties do not have significantly different pre-trends for any category of crime (p = 0.21 for violent crime, p = 0.14 for property crime, and p = 0.27 for overall crime). The results of the corresponding differential tests are reported in Appendix Table A5.

Table 9 shows that the results of our original model were robust – neither model showed any significant interaction effects. This suggests that Prop. 36 did not have a significant effect on either violent crime or property crime. These results also imply that the TSL was not important for decreasing crime. Otherwise, we would expect crime to significantly increase after relaxing the TSL, which it did not.

Table 9: Robustness check for baseline diff-in-diff

	Violent	Property	Overall
	(1)	(2)	(3)
Strict×Post	23.36	119.68	143.04
	(−7.49, 54.22)	(−113.47, 352.83)	(−98.16, 384.25)
Strict×Trend	N	N	N
Number Years	18	18	18
Number Counties	57	57	57
Observations	1,026	1,026	1,026

Note: Each column estimates equation (16) for the corresponding category of crime. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Counties are grouped based on strictness averaged over 2004-2008. The row *Strict×Trend* represents whether the corresponding regression controls for differential trends. Coefficients for economic controls, differential trend controls, and fixed effects are not shown. The sample used to estimate each regression includes all counties except Alpine County. The 95% confidence interval, calculated with robust standard errors clustered at the county-level, is given in parentheses.

We now do a robustness check on the exploration of nonlinearity. We omit Alpine County and divide the remaining counties into the four strictness buckets based on the robust definition of strictness. The second row in Table 10 below shows the relevant summary statistics. Again, counties with strictness less than the first quartile are considered super-lenient, counties between the first quartile and the median are medium-lenient, counties between the median and the third quartile are medium-strict, and counties above the third quartile are super-strict.

Table 10: Summary statistics of strictness averaged over 2004-2008

Statistic	Min	p25	Median	p75	Max	N
Strictness (all)	0.00	12.67	17.86	22.97	30.10	58
Strictness (excluding Alpine)	6.39	12.77	17.90	22.98	30.10	57

Source: California Department of Corrections and Rehabilitation (CDCR)

We test for differential trends between super-lenient counties and counties in the other strictness buckets. The results are reported in Appendix Table A6-A8. We again find that medium-lenient, medium-strict, and super-strict counties all significantly differ in pre-trends from super-lenient counties for all categories of crime. We control for these differential trends in our robustness checks, the results of which are reported in Table 11 below.

In column (1) of Table 11, we see that the results for violent crime are consistent with our baseline exploration of nonlinearity. It appears that at the 5% significance level, Prop. 36 decreased the violent crime rate in super-strict counties by 9.4%, relative to the average violent crime rate in 2012. However, like earlier, we should view this result with some caution since the joint test also shows that Prop. 36 had no significant effect on the violent crime rate in any of the counties (p = 0.18).

For property crime and overall crime, we see similar results as our baseline exploration. Neither the interaction effects nor the joint tests are significant (p = 0.53 for property crime and p = 0.49 for overall crime), which suggests that our earlier results were robust – it appears that Prop. 36 had no significant effect on property or overall crime. It is interesting to note, however, that the property crime rate represented by the upper bound of the 95% confidence interval for super-strict counties is about half that of medium-lenient and medium-strict counties: in the worst-case scenario, relative to 2012, the property crime rate in medium-lenient and medium-strict counties can increase by up to 17.5% and 21%, respectively; meanwhile, the worst-case scenario in super-strict counties corresponds to an 8.1% increase. Like our earlier results suggest, it appears that the effect of Prop. 36 depends on county strictness.

Table 11: Robustness check for exploration of nonlinearity

	Violent (1)	Property (2)	Overall (3)
MediumLenient×Post	−49.77	−16.47	−66.24
	(−116.09, 16.54)	(−713.82, 680.89)	(−783.96, 651.48)
MediumStrict×Post	−28.03	45.51	17.48
	(−87.58, 31.52)	(−724.82, 815.84)	(−780.21, 815.18)
SuperStrict×Post	−59.50**	−445.07	−504.57
	(−116.48, −2.53)	(−1,206.09, 315.95)	(−1,280.07, 270.93)
MediumLenient×Trend	Y	Y	Y
MediumStrict×Trend	Y	Y	Y
SuperStrict×Trend	Y	Y	Y
***p*-value**	**0.18**	**0.53**	**0.49**
Number Years	18	18	18
Number Counties	57	57	57
Observations	1,026	1,026	1,026

Note: Each column estimates equation (17) for the corresponding category of crime. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Counties are grouped based on strictness averaged over 2004-2008. The rows *MediumLenient×Trend*, *MediumStrict×Trend*, and *SuperStrict×Trend* represent whether the regression controls for the corresponding differential trends. The row *p-value* refers to the joint test that none of the coefficients of interest are significant. Coefficients for economic controls, differential trend controls, and fixed effects are not shown. The sample used to estimate each regression includes all counties except Alpine County. The 95% confidence interval, calculated with robust standard errors clustered at the county-level, is given in parentheses.
*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

8 Discussion

8.1 Baseline Diff-in-Diff

Our results from the baseline diff-in-diff model show that for strict counties, Prop. 36 had no significant effect on any of the crime rates we considered: the violent, property, and overall crime rate all remained the same after the reform. These results suggest that the TSL did not reduce crime as intended because if the law were actually effective, we would expect to see an increase in crime rates following the reform. The reasoning for this is that after Prop. 36 passed, it essentially forced strict counties to apply the TSL less frequently by reducing the number of crimes that count as a third strike. Thus, if these counties had successfully reduced crime before the reform by applying the TSL, then forcing them to apply it less frequently should lead to an increase in crime, i.e. Prop. 36 should cause crime rates to increase. However, our results suggest that this is not the case. Instead, strict counties saw no change in their crime rates after the reform. Based on the reasoning laid out above, this implies that the TSL was unsuccessful at reducing crime.[84]

In addition to providing insight into the effectiveness of the TSL itself, our results also inform us about the effectiveness of Prop. 36 relative to the TSL. Since we found that Prop. 36 had no significant effect on crime rates, it suggests that crime rates did not change after the reform. This means that Prop. 36 resulted in the same level of crime as the TSL and is just as effective as a sentencing policy. In other words, the same level of crime achieved under the TSL could have been achieved without sentencing non-violent third-strikers to life in prison. Thus, imposing life sentences on non-violent offenders was unnecessary, and the TSL had been excessively harsh.

Our results on the effectiveness of Prop. 36 differs from what we predict in our theoretical model. Based on our model, we expect to see an increase after the reform in both the number of strikeable and non-strikeable offenses, i.e. we expect Prop. 36 to be less effective than the TSL. Since violent and property crime are made up of only strikeable and non-strikeable offenses, our theoretical model thus predicts that both the violent and the property crime rate would increase after the reform. However, our results suggest that both crime rates remained the same.

One explanation for the inconsistency between our model and our results is that Prop. 36 could have led to a change in the composition of crimes, which our theoretical model does not account for. For example, since Prop. 36 upheld life sentencing for violent repeat offenders, first-time offenders might perceive that waiting to commit a violent crime will lead to harsher punishment. As a result, they might jump straight to committing a violent crime upfront instead of waiting and committing a few property crimes first. Likewise, those facing their second and third strike might be deterred from committing a violent crime and switch to property crimes instead. Compositional changes like these are not captured in our model and could account for the results that we see.

Another possible explanation for the discrepancy between our model and our results is that Prop. 36 might not have actually affected a potential offender's perception of delayed punishment. This could occur if a potential offender lacks full information on Prop. 36 and its consequences, in which case her perception of delayed punishment would be unaffected. Furthermore, research has shown that

---

[84] The most we can say from looking at our results is that the TSL failed to reduce crime. This implies that either the law had no effect on crime or that it had increased crime. Although we cannot infer from our results which of the two is the true effect of the TSL, we can still conclude that the TSL was ineffective as a sentencing policy.

career criminals generally do not perceive a high likelihood of apprehension.[85] If that were the case, then changing the punishment for repeat offenders, even if they were fully informed, would not affect their decision to commit any crimes. Since repeat offenders believe that their getting caught is very unlikely, their decision to offend would depend very little on the punishment incurred if caught, and even less so on future punishment. Thus, relaxing the TSL in the way that Prop. 36 did would have no impact on crime rates.

In addition, our theoretical model of crime rests on the assumption that individuals are rational actors given the information received. Our model does not account for factors like age, impulsivity, mental illness, or antisocial personality disorder, which have all been found to be correlated with an individual's decision to offend.[86] Mental illness, in particular, is particularly relevant to TSL sentencing. Michael Romano, director of the Three Strikes Project at Stanford Law School, has found from his experience that everyone who is sentenced to life in prison for a non-serious or non-violent offense "suffers from some kind of mental illness or impairment."[87] Research has also shown that "a fair number" of individuals offend while under the influence of drugs and/or alcohol, which would have hindered their ability to reason at the time of their offense.[88] All of these are potential factors that could undermine our model's assumption that offenders are rational beings. As a result, if they are unable to rationally evaluate the reform and its implications, then Prop. 36 would have no effect on their behavior, as suggested by our results.

8.2 Exploration of Nonlinearity

Our baseline diff-in-diff model rests on the assumption that crime varies linearly with strictness. In order to examine this assumption, we ran an exploration of nonlinearity by dividing the counties into four buckets of strictness instead of two. Our results suggest that the effect of Prop. 36 on violent crime depends on how strictly a county applied the TSL prior to the reform. While Prop. 36 had no significant effect on the violent crime rate in medium-strict and medium-lenient counties, it significantly decreased the violent crime rate by 11.5% in super-strict counties (*p*

---

[85] George Bridges and James Stone, "Effects of Criminal Punishment on Perceived Threat of Punishment: Toward an Understanding of Specific Deterrence," *Journal on Research in Crime and Delinquency* 23, no. 3 (August 1, 1986): 207–39, https://doi.org/10.1177/0022427886023003002.

[86] Lee Ellis, Kevin M. Beaver, and John Wright, *Handbook of Crime Correlates* (Academic Press, 2009).

[87] Brent Staples, "California Horror Stories and the 3-Strikes Law," *The New York Times*, November 24, 2012, sec. Opinion, https://www.nytimes.com/2012/11/25/opinion/sunday/california-horror-stories-and-the-3-strikes-law.html.

[88] Meg Chapman et al., "ADAM II: 2009 Annual Report," *Office of National Drug Control Policy*, June 2010, 182.

= 0.021), relative to the average violent crime rate in 2012. It is important to note, however, that the joint test for violent crime indicates that we should view these results with some caution. Meanwhile, our results for property and overall crime are inconclusive; we do not have enough statistical power to determine whether Prop. 36 had any effect on either the property or the overall crime rate.

Like the baseline diff-in-diff, our exploration of nonlinearity also suggests that the TSL was ineffective at reducing violent crime. The reasoning is the same as described earlier in this section. Since medium-lenient, medium-strict, and super-strict counties all applied the TSL more frequently than super-lenient counties, they were all forced to reduce their respective frequencies after Prop. 36. Thus, if they had successfully reduced violent crime before the reform by applying the TSL as frequently as they did, then forcing them to apply it less should lead to an increase in their violent crime rate. However, no group experienced a significant increase in its violent crime rate after the reform. This suggests that before Prop. 36, the TSL was not effective at reducing violent crime in any county, regardless of how frequently it was applied. This is consistent with what we found using the baseline diff-in-diff model.

What is surprising, however, is that the effectiveness of Prop. 36 as a sentencing policy appears to depend on a county's strictness level. Since Prop. 36 decreased the violent crime rate in super-strict counties while not affecting the violent crime rate in other counties, it suggests that there exists a strictness-threshold at which Prop. 36 actually becomes more effective than the existing TSL at reducing violent crime. Counties that lie above the strictness-threshold (i.e. super-strict counties) will be better off applying Prop. 36 rather than their super-strict version of the TSL. Meanwhile, in counties that fall below this strictness threshold (i.e. medium-strict and medium-lenient counties), our results suggest that Prop. 36 is just as effective as the more lenient versions of the TSL that they had been using.

Furthermore, our results for super-strict counties suggest that Prop. 36 was actually more effective than the TSL itself at reducing violent crime. The reasoning is as follows. Since super-strict counties saw their violent crime rate decrease after the reform, it implies that Prop. 36 is more effective at reducing violent crime than the super-strict version of the TSL they had been implementing before. Since super-strict counties are those that apply the TSL most frequently, they can be thought of as counties that had fully implemented the TSL. Thus, Prop. 36 is more effective than the TSL (as intended) when it comes to reducing violent crime – restricting mandatory life sentences to only strikeable offenses leads to less violent crime than broadly applying such a severe punishment to *all* offenses. If super-strict counties had implemented Prop. 36 instead, they would have achieved a lower violent crime rate. Thus, closely adhering to the TSL is actually counterproductive to fighting violent crime – it leads to more violent crime than is necessary.

One potential explanation for why Prop. 36 is more effective than the TSL at reducing violent crime is that Prop. 36 could have made the punishment for violent crime appear more severe, thus deterring potential violent offenders. For example, Prop. 36 was framed on the ballot as a way of doubling down on "dangerous career criminals" and preventing them from being released early from prison.[89] This could lead potential offenders to think (at least for a short period of time) that committing a violent crime under Prop. 36 leads to a higher likelihood of conviction and a more severe sentence if convicted. Since deterrence is based on an individual's perception of the certainty and severity of punishment, this change in perception would deter potential offenders from committing a violent crime. Deterrence would then lead to the decrease in the violent crime rate that we see in our results.

In addition to deterring potential violent offenders, Prop. 36 also could have incentivized them to switch to non-strikeable offenses. Our theoretical model of crime shows that individuals facing their last strike after the reform would receive relatively more utility from committing a non-strikeable offense than a strikeable one (at least for a short period of time). Thus, if these individuals were to commit another crime, they could be incentivized to substitute away from violent crimes.

Our results are less precise for property and overall crime. We do not have enough statistical power to determine whether Prop. 36 had any effect on the property and overall crime rate in each type of county. However, it is interesting to note that the worst-case scenario is much less serious for super-strict counties than for medium-strict and medium-lenient counties – there seems to be a ceiling to the detrimental effect that Prop. 36 can have on counties that had applied the TSL most frequently. This suggests that the effect of Prop. 36 depends to some degree on a county's strictness level.

8.3 Policy Implications

Overall, our results suggest that sentencing offenders to life in prison for a non-violent crime is a poor sentencing policy. Doing so does not reduce crime. Instead, our results suggest that a better policy is one that embraces proportional sentencing such that the punishment increases with the gravity of the crime. Prop. 36 takes the TSL in the direction of proportional sentencing by giving non-violent repeat offenders double the typical sentence rather than a mandatory life sentence. In doing so, Prop. 36 actually leads to fewer violent crimes than the TSL. This is especially interesting since the main motivation for the TSL's extreme sentencing framework was to reduce serious and violent crime. However, our results suggest that the most severe element of the TSL—punishing non-strikeable crimes with life in prison—was, in fact, too severe.

---

[89] Voter Information Guide for 2012, General Election (2012). http://repository.uchastings.edu/ca_ballot_props/1320

The assumption that motivated the TSL's third-strike provision was that repeat offenders are not deterred by the length of incarceration and will continue to engage in crime regardless. Thus, it was concluded that the only "deterrent value" of incarceration was to signal the "reality and inevitability of punishment," which the TSL attempted to achieve through mandatory life sentences.[90] However, our results suggest that the specific length of incarceration does matter – proportional sentencing is crucial to deterring potential offenders, especially potential violent offenders. When there is no differential punishment between violent and non-violent crimes, it may lead to an increase in violent crimes since the severity of punishment is no longer a deterring factor. For those who have already been convicted of two strikeable offenses, they face no cost in becoming violent. If they knew that they would receive a severe sentence regardless of the crime that is committed, they would have no incentive to risk their last strike with a petty crime. Proportionality in sentencing is thus necessary to prevent violent crimes – it encourages those who have a disposition towards criminal activities to commit lower-level offenses.

Even if we took a more cautious approach to interpreting our results and concluded that Prop. 36 was not any more effective as a sentencing policy than the TSL, the former is still preferable to the latter from a fiscal standpoint. According to the nonpartisan California Legislative Analyst's Office, the TSL cost about $500 million per year to implement. These expenses would have only escalated over time as the inmate population aged and required additional medical expenses. This is especially true for third-striker inmates, almost half of whom had been locked up for life over a non-strikeable offense.[91] In 1980, corrections expenditure only took up 3% of the state's General Fund. By 2010, it accounted for more than 10% of the budget.[92]

The increasing cost of the TSL inevitably shifted resources away from other services that had been funded by the state budget. In 2008, California faced a $16 billion budget deficit and as a result, was forced to cut spending in many areas, including education, medical benefits, and welfare benefits. State prisons were also forced to release some of their inmates early due to lack of funding.[93] The opportunity cost of implementing the TSL is thus extremely high and, as such, deserves serious consideration.

Such high costs may be justified if the TSL had reduced crime and saved taxpayer money. However, our results suggest that not only did the TSL fail to reduce crime but also that the same level of crime could have been achieved with a

---

[90] Ardaiz, ibid.

[91] Brown and Jolivette, ibid.

[92] Lofstrom, Bird, and Martin, ibid.

[93] Chen, "Impacts of…," ibid.

more efficient sentencing policy. Eighteen months after the reform, the Three Strikes Project reported that Prop. 36 had already cut California's prison costs by over $30 million.[94] California's Legislative Analyst estimates that Prop. 36 can eventually save the state up to $90 million annually.[95] These savings mostly stem from the reduced sentences for non-violent third-strikers as well as the resentencing of eligible third-strikers already in prison before the reform. Prop. 36 would thus free up hundreds of millions of dollars for programs that could be more effective at reducing crime than simply locking up repeat offenders for low-level offenses. Potential areas to invest in include education, job creation, inmate rehabilitation, and mental health services.

In addition to the substantial financial cost that it incurs, the TSL also imposed serious non-monetary costs on the human and societal level. When the TSL sentences individuals to prison for life, it results in many unintended and serious consequences. One such consequence is that it leads to more single-parent families.[96] This is detrimental because children raised in single-parent households are often adversely affected in terms of educational attainment and income level and, as a result, may be less likely to secure a job in the future.[97] If these children have trouble finding legitimate earnings opportunities later on, they might turn to criminal activities for income instead, thus creating a perpetual cycle of crime. It is also worth noting that the TSL targets vulnerable populations and disproportionately locks up individuals from marginalized communities. Most non-violent third-strikers suffer from mental illnesses and African American men, who constitute about 3% of California's population, represent about 44% of third-striker inmates.[98]

---

94 Stanford Law School Three Strikes Project and NAACP Legal Defense and Education Fund, ibid.

95 "Proposition 36: Three Strikes Law. Sentencing for Repeat Felony Offenders. Initiative Statute.," Legislative Analyst's Office, July 18, 2012, https://lao.ca.gov/ballot/2012/36_11_2012.aspx.

96 MARK GIUS, "THE UNINTENDED CONSEQUENCES OF THREE STRIKES LAWS: The Impact of Long Term Imprisonments on Family Structures," *International Journal of Sociology of the Family* 43, no. 1/2 (2017): 17–24.

97 Daniel P. Mueller and Philip W. Cooper, "Children of Single Parent Families: How They Fare as Young Adults," *Family Relations* 35, no. 1 (1986): 169–76, https://doi.org/10.2307/584296; Douglas B. Downey, "The School Performance of Children From Single-Mother and Single-Father Families," *Journal of Family Issues* 15, no. 1 (March 1, 1994): 129–47, https://doi.org/10.1177/019251394015001006;David T. Ellwood and Christopher Jencks, "The Spread of Single-Parent Families in the United States Since 1960," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, February 26, 2004), https://doi.org/10.2139/ssrn.517662.

98 Chen, "Impacts of…," ibid.

9 Conclusion

This paper examines the effect of Prop. 36 on various crime rates relative to the TSL. We use the natural experiment provided by the passage of Prop. 36 to examine whether the TSL had been too severe as a sentencing law. Specifically, we consider whether imposing life sentences on non-violent offenders had been necessary.

We find that for violent crime, the effect of Prop. 36 depends on how strictly a county had applied the TSL prior to the reform. Our results suggest that counties that had closely adhered to the TSL experienced a significant decrease in their violent crime rate following the reform, while counties that had implemented a more lenient version of the TSL were unaffected. This suggests that adhering closely to the TSL is actually detrimental to reducing violent crime and that Prop. 36 is more effective at reducing violent crime than the TSL itself. Thus, our results demonstrate that the TSL had been too severe. By providing for an extremely harsh punishment on non-violent third-strikers, the TSL had inadvertently led to more violent crime than was necessary.

Additionally, we also find that Prop. 36 had no significant effect on either property crime or overall crime. This indicates that the same level of property and overall crime achieved under the TSL could have been achieved through Prop. 36. In other words, imposing a life sentence on non-violent offenders was unnecessary and excessively harsh.

Finally, because no crime rates significantly increased after the reform, it implies that the TSL had been ineffective as a sentencing policy in general. If it had been successful at reducing crime, then relaxing the law should lead to an increase in crime. However, our results suggest that no such increase occurred.

Mandatory life sentences may be a fair punishment in response to heinous crimes, like murder or rape, but when it is used to punish an individual for a petty crime, like stealing a slice of pizza, it becomes much harder to justify. In order to protect the integrity of our criminal justice system and the public's perception of its legitimacy, it is imperative that we have a fair criminal sentencing policy. We find that Prop. 36 is a step in the right direction. Whether it is the most optimal sentencing policy is beyond the scope of this paper. Further research is needed to determine the best approach to reducing crime, one that balances fiscal costs and social repercussions.

References

Arango, Tim. "In California, Criminal Justice Reform Offers a Lesson for the Nation (Published 2019)." *The New York Times*, January 21, 2019, sec. U.S. https://www.nytimes.com/2019/01/21/us/california-incarceration-reduction-penalties.html.

Ardaiz, James A. "California's Three Strikes Law: History, Expectations, Consequences." *McGeorge Law Review* 32 (2000). https://heinonline.org/HOL/Page?handle=hein.journals/mcglr32&id=11&div=&collection=.

Austin, James, John Clark, Patricia Hardyman, and Alan Henry. "Three Strikes and You're Out: The Implementation and Impart of Strike Laws," March 6, 2000.

Becker, Gary S. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76, no. 2 (1968): 169–217.

Bridges, George, and James Stone. "Effects of Criminal Punishment on Perceived Threat of Punishment: Toward an Understanding of Specific Deterrence." *Journal on Research in Crime and Delinquency* 23, no. 3 (August 1, 1986): 207–39. https://doi.org/10.1177/0022427886023003002.

Brown, Brian, and Greg Jolivette. "A Primer: Three Strikes - The Impact After More Than a Decade." Legislative Analyst's Office, October 2005. https://lao.ca.gov/2005/3_strikes/3_strikes_102005.htm.

Ballotpedia. "California Proposition 36, Changes to Three Strikes Sentencing Initiative (2012)." Accessed April 19, 2021. https://ballotpedia.org/California_Proposition_36,_Changes_to_Three_Strikes_Sentencing_Initiative_(2012).

"California Proposition 184, Three Strikes Sentencing Initiative (1994) - Ballotpedia." Accessed April 19, 2021. https://ballotpedia.org/California_Proposition_184,_Three_Strikes_Sentencing_Initiative_(1994).

Chapman, Meg, Sara Collins, Chris Flygare, Dana Hunt, Sarah Kuck-Jalbert, Ryan Kling, Richard Kulka, Caroline Nobo, and William Rhodes. "ADAM II: 2009 Annual Report." *Office of National Drug Control Policy*, June 2010, 182.

Chen, Elsa. "The Liberation Hypothesis and Racial and Ethnic Disparities in the Application of California's Three Strikes Law." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 2008. https://papers.ssrn.com/abstract=2189506.

Chen, Elsa Y. “Impacts of ‘Three Strikes and You’re Out’ on Crime Trends in California and Throughout the United States.” *Journal of Contemporary Criminal Justice* 24, no. 4 (November 1, 2008): 345–70. https://doi.org/10.1177/1043986208319456.

Downey, Douglas B. “The School Performance of Children From Single-Mother and Single-Father Families.” *Journal of Family Issues* 15, no. 1 (March 1, 1994): 129–47. https://doi.org/10.1177/019251394015001006.

Ehrlich, Isaac. “Crime, Punishment, and the Market for Offenses.” *Journal of Economic Perspectives* 10, no. 1 (February 1, 1996): 43–67. https://doi.org/10.1257/jep.10.1.43.

Ellis, Lee, Kevin M. Beaver, and John Wright. *Handbook of Crime Correlates*. Academic Press, 2009.

Ellwood, David T., and Christopher Jencks. “The Spread of Single-Parent Families in the United States Since 1960.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, February 26, 2004. https://doi.org/10.2139/ssrn.517662.

Gabbard, Ashley, Krista Christian, Shelley Buttler, John Yessen, Michael Keeling, and Yvonne Lawrence. “Offender Data Points,” 2019, 177.

GIUS, MARK. “THE UNINTENDED CONSEQUENCES OF THREE STRIKES LAWS: The Impact of Long Term Imprisonments on Family Structures.” *International Journal of Sociology of the Family* 43, no. 1/2 (2017): 17–24.

Golaszewski, Paul. “A Status Report: Reducing Prison Overcrowding in California.” Legislative Analyst’s Office, August 5, 2011. https://lao.ca.gov/reports/2011/crim/overcrowding_080511.aspx.

Helland, Eric, and Alexander Tabarrok. “Does Three Strikes Deter? A Nonparametric Estimation.” *Journal of Human Resources* XLII, no. 2 (March 31, 2007): 309–30. https://doi.org/10.3368/jhr.XLII.2.309.

Iyengar, Radha. “I’d Rather Be Hanged for a Sheep than a Lamb: The Unintended Consequences of ‘Three-Strikes’ Laws.” National Bureau of Economic Research, February 7, 2008. https://doi.org/10.3386/w13784.

Jackson, Pamela Irving, and Leo Carroll. “Race and the War on Crime: The Sociopolitical Determinants of Municipal Police Expenditures in 90 Non-Southern U.S. Cities.” *American Sociological Review* 46, no. 3 (1981): 290–305. https://doi.org/10.2307/2095061.

Lofstrom, Magnus, Mia Bird, and Brandon Martin. “California’s Historic Corrections Reforms.” *Public Policy Institute of California* (blog), September 2016. https://www.ppic.org/publication/californias-historic-corrections-reforms/.

Lofstrom, Magnus, and Brandon Martin. “Crime Trends in California.” *Public Policy Institute of California* (blog), February 2021. https://www.ppic.org/publication/crime-trends-in-california/.

Marvell, Thomas B., and Carlisle E. Moody. “The Lethal Effects of Three-Strikes Laws.” *The Journal of Legal Studies* 30, no. 1 (January 1, 2001): 89–106. https://doi.org/10.1086/468112.

McPheters, Lee R., and William B. Stronge. “Law Enforcement Expenditures and Urban Crime.” *National Tax Journal* 27, no. 4 (December 1, 1974): 633–44. https://doi.org/10.1086/NTJ41861994.

Misczynski, Dean. “Corrections Realignment: One Year Later.” *Public Policy Institute of California*, August 2012, 38.

Mueller, Daniel P., and Philip W. Cooper. “Children of Single Parent Families: How They Fare as Young Adults.” *Family Relations* 35, no. 1 (1986): 169–76. https://doi.org/10.2307/584296.

Pishko, Jessica. “In Liberal California, A Crusader Against Criminal Justice Reform.” The Appeal, August 19, 2019. https://theappeal.org/in-liberal-california-a-crusader-against-criminal-justice-reform/.

Press, The Associated. “11th-Hour Attempt Overturns Prop 66.” The Santa Clara, November 3, 2004. https://www.thesantaclara.org/blog/11th-hour-attempt-overturns-prop-66.

———. “25 Years for a Slice of Pizza.” *The New York Times*, March 5, 1995, sec. U.S. https://www.nytimes.com/1995/03/05/us/25-years-for-a-slice-of-pizza.html.

Legislative Analyst’s Office. “Proposition 36: Three Strikes Law. Sentencing for Repeat Felony Offenders. Initiative Statute.,” July 18, 2012. https://lao.ca.gov/ballot/2012/36_11_2012.aspx.

Romano, Michael. “Divining the Spirit of California’s Three Strikes Law.” *Federal Sentencing Reporter* 22, no. 3 (2010): 171–75. https://doi.org/10.1525/fsr.2010.22.3.171.

Shepherd, Joanna M. “Fear of the First Strike: The Full Deterrent Effect of California’s Two- and Three-Strikes Legislation.” *The Journal of Legal Studies* 31, no. 1 (January 1, 2002): 159–201. https://doi.org/10.1086/324660.

Skelton, George. “A Father’s Crusade Born From Pain.” Los Angeles Times, December 9, 1993. https://www.latimes.com/archives/la-xpm-1993-12-09-mn-65402-story.html.

Stanford Law School Three Strikes Project, and NAACP Legal Defense and Education Fund. “Proposition 36 Progress Report: Over 1,500 Prisoners Released; Historically Low Recidivism Rate,” April 2014. https://law.stanford.edu/publications/proposition-36-progress-report-over-1500-prisoners-released-historically-low-recidivism-rate/.

Staples, Brent. "California Horror Stories and the 3-Strikes Law." *The New York Times*, November 24, 2012, sec. Opinion. https://www.nytimes.com/2012/11/25/opinion/sunday/california-horror-stories-and-the-3-strikes-law.html.

Ballotpedia. "Text of California Proposition 36 (November 2012)." Accessed December 21, 2020. https://ballotpedia.org/Text_of_California_Proposition_36_(November_2012).

Tomlinson, Kelli D. "An Examination of Deterrence Theory: Where Do We Stand." *Federal Probation* 80, no. 3 (2016): 33–38.

United States. Bureau of Justice Statistics. National Corrections Reporting Program, [United States], 2000-2017. Inter-university Consortium for Political and Social Research [distribu-tor], 2020-11-19. https://doi.org/10.3886/ICPSR37608.v1

Vitiello, Michael. "Three Strikes: Can We Return to Rationality?" *The Journal of Criminal Law and Criminology (1973-)* 87, no. 2 (1997): 395. https://doi.org/10.2307/1143951.

"Voter Information Guide for 2012, General Election," 2012, 145.

Weisberg, Robert. "The Wild West of Sentencing Reform: Lessons from California." *Crime and Justice* 48 (February 8, 2019): 35–77. https://doi.org/10.1086/701714.

Worrall, John L. "The Effect of Three-Strikes Legislation on Serious Crime in California." *Journal of Criminal Justice* 32, no. 4 (July 1, 2004): 283–96. https://doi.org/10.1016/j.jcrimjus.2004.04.001.

Appendices

Table A1: Parallel trends test

	Violent	Property	Overall
	(1)	(2)	(3)
Strict×2000	−103.21	−36.51	−139.72
	(66.37)	(446.48)	(479.21)
Strict×2001	−92.57	−749.03*	−841.60*
	(66.66)	(449.34)	(503.02)
Strict×2002	−50.32	−531.10	−581.42
	(67.82)	(485.78)	(542.61)
Strict×2003	−127.10	−710.81	−837.91
	(78.78)	(514.33)	(580.99)
Strict×2004	−74.06	−794.80	−868.86
	(66.04)	(540.07)	(586.54)
Strict×2005	−63.55	−249.58	−313.13
	(100.31)	(381.28)	(465.68)
Strict×2006	−93.46	−293.45	−386.92
	(75.06)	(358.97)	(407.60)
Strict×2007	−36.08	−184.31	−220.39
	(49.33)	(433.80)	(464.25)
Strict×2008	−147.02**	−296.09	−443.12
	(69.13)	(406.91)	(450.76)
Strict×2009	−101.58	−301.98	−403.56
	(65.10)	(428.56)	(481.15)
Strict×2010	24.23	−232.48	−208.25
	(87.11)	(508.44)	(571.51)
Strict×2011	−110.68	−450.45	−561.14
	(69.21)	(444.84)	(503.77)
***p*-value**	**0.034**	**0.38**	**0.42**
Number Years	18	18	18
Number Counties	57	57	57
Observations	1,026	1,026	1,026

Note: The regression here tests for differential trends between strict and lenient counties. Counties are grouped based on strictness averaged over 2004-2012. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Each row corresponds to the coefficient on the interaction term between *Strict* and a dummy variable for every year in the pre-period before 2012. The row *p-value* refers to the joint test that none of the pre-period interaction effects are significant. Coefficients for post-period interaction terms, economic controls, and fixed effects are not shown. The sample used to estimate each regression includes all counties except Alpine County. Robust standard errors clustered at the county-level are given in parentheses.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table A2: Parallel trends test: medium-lenient vs. super-lenient

	Violent	Property	Overall
	(1)	(2)	(3)
MediumLenient×2000	−8.07	105.37	97.29
	(90.49)	(574.21)	(620.07)
MediumLenient×2001	78.15	252.81	330.96
	(95.71)	(708.87)	(793.37)
MediumLenient×2002	1.55	−240.05	−238.50
	(92.22)	(715.82)	(794.31)
MediumLenient×2003	35.41	439.64	475.06
	(111.45)	(758.45)	(857.41)
MediumLenient×2004	131.39	478.27	609.67
	(102.30)	(780.88)	(849.95)
MediumLenient×2005	86.64	573.65	660.29
	(122.24)	(518.46)	(609.48)
MediumLenient×2006	70.20	467.72	537.92
	(104.22)	(530.17)	(599.52)
MediumLenient×2007	131.94*	509.36	641.30
	(75.88)	(623.12)	(676.13)
MediumLenient×2008	18.40	−31.92	−13.51
	(113.82)	(558.58)	(621.30)
MediumLenient×2009	120.97	238.15	359.12
	(95.89)	(588.32)	(669.20)
MediumLenient×2010	173.29	524.36	697.65
	(109.46)	(690.04)	(775.36)
MediumLenient×2011	130.36	410.07	540.43
	(102.36)	(637.95)	(727.67)
***p*-value**	**< 0.001**	**< 0.001**	**0.002**
Number Years	18	18	18
Number Counties	29	29	29
Observations	522	522	522

Note: The regression here tests for differential trends between medium-lenient and super-lenient counties. Counties are grouped based on strictness averaged over 2004-2012. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Each row corresponds to the coefficient on the interaction term between *MediumLenient* and a dummy variable for every year in the pre-period before 2012. The row *p-value* refers to the joint test that none of the pre-period interaction effects are significant. Coefficients for post-period interaction terms, economic controls, and fixed effects are not shown. Robust standard errors clustered at the county-level are given in parentheses.

*p < 0.1, **p < 0.05, ***p < 0.01

Table A3: Parallel trends test: medium-strict vs. super-lenient

	Violent	Property	Overall
	(1)	(2)	(3)
MediumStrict×2000	−39.23	101.79	62.57
	(86.94)	(668.25)	(712.50)
MediumStrict×2001	−8.68	−790.65	−799.33
	(100.29)	(684.72)	(772.49)
MediumStrict×2002	4.64	−508.25	−503.62
	(114.64)	(814.81)	(920.80)
MediumStrict×2003	−85.31	−581.41	−666.71
	(114.76)	(752.46)	(849.40)
MediumStrict×2004	−1.03	−760.03	−761.06
	(89.73)	(769.83)	(830.55)
MediumStrict×2005	56.87	201.96	258.84
	(156.89)	(607.28)	(750.91)
MediumStrict×2006	21.86	−95.08	−73.22
	(102.82)	(545.27)	(616.61)
MediumStrict×2007	57.87	−182.50	−124.63
	(59.95)	(622.69)	(660.42)
MediumStrict×2008	−120.38	−430.63	−551.01
	(112.45)	(657.72)	(722.34)
MediumStrict×2009	−2.63	−203.07	−205.69
	(93.10)	(642.72)	(722.57)
MediumStrict×2010	190.27	−29.14	161.13
	(127.70)	(771.92)	(874.35)
MediumStrict×2011	−34.58	−575.16	−609.74
	(100.72)	(679.04)	(768.26)
***p*-value**	**< 0.001**	**0.004**	**0.005**
Number Years	18	18	18
Number Counties	28	28	28
Observations	504	504	504

Note: The regression here tests for differential trends between medium-strict and super-lenient counties. Counties are grouped based on strictness averaged over 2004-2012. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Each row corresponds to the coefficient on the interaction term between *MediumStrict* and a dummy variable for every year in the pre-period before 2012. The row *p-value* refers to the joint test that none of the pre-period interaction effects are significant. Coefficients for post-period interaction terms, economic controls, and fixed effects are not shown. Robust standard errors clustered at the county-level are given in parentheses.

Table A4: Parallel trends test: super-strict vs. super-lenient

	Violent	Property	Overall
	(1)	(2)	(3)
SuperStrict×2000	−162.89*	35.84	−127.05
	(98.54)	(749.07)	(789.68)
SuperStrict×2001	−82.76	−350.68	−433.44
	(97.35)	(742.36)	(824.72)
SuperStrict×2002	−94.58	−746.50	−841.08
	(87.01)	(727.43)	(799.91)
SuperStrict×2003	−124.22	−353.83	−478.05
	(111.35)	(819.88)	(919.20)
SuperStrict×2004	−2.11	−312.32	−314.43
	(85.53)	(800.21)	(856.15)
SuperStrict×2005	−86.58	−86.81	−173.39
	(133.79)	(577.41)	(687.42)
SuperStrict×2006	−130.53	6.43	−124.10
	(110.17)	(565.22)	(642.63)
SuperStrict×2007	10.99	340.58	351.57
	(60.49)	(611.47)	(647.32)
SuperStrict×2008	−149.65	−184.98	−334.63
	(103.83)	(531.73)	(582.48)
SuperStrict×2009	−71.95	−154.62	−226.58
	(94.12)	(608.38)	(688.96)
SuperStrict×2010	41.38	111.87	153.26
	(114.26)	(683.76)	(765.45)
SuperStrict×2011	−49.35	95.59	46.24
	(98.10)	(609.69)	(697.67)
***p*-value**	**0.039**	**0.034**	**0.020**
Number Years	18	18	18
Number Counties	28	28	28
Observations	504	504	504

Note: The regression here tests for differential trends between super-strict and super-lenient counties. Counties are grouped based on strictness averaged over 2004-2012. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Each row corresponds to the coefficient on the interaction term between *SuperStrict* and a dummy variable for every year in the pre-period before 2012. The row *p-value* refers to the joint test that none of the pre-period interaction effects are significant. Coefficients for post-period interaction terms, economic controls, and fixed effects are not shown. Robust standard errors clustered at the county-level are given in parentheses.

*p < 0.1, **p < 0.05, ***p < 0.01

Table A5: Parallel trends test: robustness check for baseline diff-in-diff

	Violent	Property	Overall
	(1)	(2)	(3)
Strict×2000	−99.83	6.21	−93.62
	(65.77)	(444.04)	(476.64)
Strict×2001	−135.62**	−759.26*	−894.88*
	(65.30)	(452.16)	(504.79)
Strict×2002	−104.20	−754.20	−858.40
	(66.73)	(483.92)	(539.58)
Strict×2003	−163.12**	−685.69	−848.81
	(77.56)	(517.38)	(583.49)
Strict×2004	−86.15	−680.95	−767.10
	(66.40)	(542.92)	(589.41)
Strict×2005	−130.97	−330.15	−461.12
	(98.17)	(378.82)	(460.73)
Strict×2006	−139.43*	−267.08	−406.51
	(73.56)	(358.34)	(406.27)
Strict×2007	−63.57	−126.99	−190.56
	(49.27)	(432.53)	(463.37)
Strict×2008	−169.02**	−437.71	−606.73
	(68.75)	(401.57)	(444.20)
Strict×2009	−127.60**	−334.12	−461.72
	(64.65)	(428.46)	(480.87)
Strict×2010	−46.15	−235.14	−281.29
	(86.71)	(509.16)	(571.74)
Strict×2011	−122.22*	−281.64	−403.86
	(69.29)	(449.38)	(509.05)
***p*-value**	**0.21**	**0.14**	**0.27**
Number Years	18	18	18
Number Counties	57	57	57
Observations	1,026	1,026	1,026

Note: The regression here tests for differential trends between strict and lenient counties. Counties are grouped based on strictness averaged over 2004-2008. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Each row corresponds to the coefficient on the interaction term between *Strict* and a dummy variable for every year in the pre-period before 2012. The row *p-value* refers to the joint test that none of the pre-period interaction effects are significant. Coefficients for post-period interaction terms, economic controls, and fixed effects are not shown. The sample used to estimate each regression includes all counties except Alpine County. Robust standard errors clustered at the county-level are given in parentheses.

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table A6: Parallel trends test: robustness check for medium-lenient vs. super-lenient

	Violent	Property	Overall
	(1)	(2)	(3)
MediumLenient×2000	−7.44	37.23	29.79
	(80.13)	(551.12)	(585.87)
MediumLenient*2001	128.51	269.14	397.65
	(93.87)	(704.43)	(788.28)
MediumLenient*2002	59.06	−29.01	30.05
	(103.82)	(766.26)	(857.90)
MediumLenient*2003	78.54	430.76	509.30
	(109.99)	(754.05)	(852.58)
MediumLenient*2004	151.67	387.25	538.92
	(105.15)	(774.26)	(846.67)
MediumLenient*2005	160.91	666.35	827.26
	(127.69)	(527.20)	(625.46)
MediumLenient*2006	122.67	454.56	577.23
	(101.26)	(523.17)	(591.93)
MediumLenient*2007	165.40**	467.05	632.46
	(79.34)	(624.06)	(680.29)
MediumLenient*2008	46.38	111.81	158.19
	(112.00)	(554.84)	(613.79)
MediumLenient*2009	155.19*	281.06	436.25
	(90.60)	(594.68)	(672.32)
MediumLenient*2010	250.27**	545.36	795.64
	(113.72)	(742.19)	(834.20)
MediumLenient*2011	150.79	263.38	414.17
	(103.39)	(668.81)	(762.16)
Joint *p*-value	**< 0.001**	**< 0.001**	**< 0.001**
Number Years	18	18	18
Number Counties	28	28	28
Observations	504	504	504

Note: The regression here tests for differential trends between medium-lenient and super-lenient counties. Counties are grouped based on strictness averaged over 2004-2008. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Each row corresponds to the coefficient on the interaction term between *MediumLenient* and a dummy variable for every year in the pre-period before 2012. The row *p-value* refers to the joint test that none of the pre-period interaction effects are significant. Coefficients for post-period interaction terms, economic controls, and fixed effects are not shown. Robust standard errors clustered at the county-level are given in parentheses.

*p < 0.1, **p < 0.05, ***p < 0.01

Table A7: Parallel trends test: robustness check for medium-strict vs. super-lenient

	Violent	Property	Overall
	(1)	(2)	(3)
MediumStrict×2000	−7.44	37.23	29.79
	(80.13)	(551.12)	(585.87)
MediumStrict×2001	128.51	269.14	397.65
	(93.87)	(704.43)	(788.28)
MediumStrict×2002	59.06	−29.01	30.05
	(103.82)	(766.26)	(857.90)
MediumStrict×2003	78.54	430.76	509.30
	(109.99)	(754.05)	(852.58)
MediumStrict×2004	151.67	387.25	538.92
	(105.15)	(774.26)	(846.67)
MediumStrict×2005	160.91	666.35	827.26
	(127.69)	(527.20)	(625.46)
MediumStrict×2006	122.67	454.56	577.23
	(101.26)	(523.17)	(591.93)
MediumStrict×2007	165.40**	467.05	632.46
	(79.34)	(624.06)	(680.29)
MediumStrict×2008	46.38	111.81	158.19
	(112.00)	(554.84)	(613.79)
MediumStrict×2009	155.19*	281.06	436.25
	(90.60)	(594.68)	(672.32)
MediumStrict×2010	250.27**	545.36	795.64
	(113.72)	(742.19)	(834.20)
MediumStrict×2011	150.79	263.38	414.17
	(103.39)	(668.81)	(762.16)
Joint *p*-value	**< 0.001**	**0.008**	**0.021**
Number Years	18	18	18
Number Counties	28	28	28
Observations	504	504	504

Note: The regression here tests for differential trends between medium-strict and super-lenient counties. Counties are grouped based on strictness averaged over 2004-2008. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Each row corresponds to the coefficient on the interaction term between *MediumStrict* and a dummy variable for every year in the pre-period before 2012. The row *p-value* refers to the joint test that none of the pre-period interaction effects are significant. Coefficients for post-period interaction terms, economic controls, and fixed effects are not shown. Robust standard errors clustered at the county-level are given in parentheses.

*p < 0.1, **p < 0.05, ***p < 0.01

Table A8: Parallel trends test: robustness check for super-strict vs. super-lenient

	Violent	Property	Overall
	(1)	(2)	(3)
SuperStrict×2000	−7.44	37.23	29.79
	(80.13)	(551.12)	(585.87)
SuperStrict×2001	128.51	269.14	397.65
	(93.87)	(704.43)	(788.28)
SuperStrict×2002	59.06	−29.01	30.05
	(103.82)	(766.26)	(857.90)
SuperStrict×2003	78.54	430.76	509.30
	(109.99)	(754.05)	(852.58)
SuperStrict×2004	151.67	387.25	538.92
	(105.15)	(774.26)	(846.67)
SuperStrict×2005	160.91	666.35	827.26
	(127.69)	(527.20)	(625.46)
SuperStrict×2006	122.67	454.56	577.23
	(101.26)	(523.17)	(591.93)
SuperStrict×2007	165.40**	467.05	632.46
	(79.34)	(624.06)	(680.29)
SuperStrict×2008	46.38	111.81	158.19
	(112.00)	(554.84)	(613.79)
SuperStrict×2009	155.19*	281.06	436.25
	(90.60)	(594.68)	(672.32)
SuperStrict×2010	250.27**	545.36	795.64
	(113.72)	(742.19)	(834.20)
SuperStrict×2011	150.79	263.38	414.17
	(103.39)	(668.81)	(762.16)
Joint *p*-value	**0.045**	**0.012**	**0.014**
Number Years	18	18	18
Number Counties	28	28	28
Observations	504	504	504

Note: The regression here tests for differential trends between super-strict and super-lenient counties. Counties are grouped based on strictness averaged over 2004-2008. The outcome variable is the corresponding crime rate, which is given as number of offenses per 100k people. Each row corresponds to the coefficient on the interaction term between *SuperStrict* and a dummy variable for every year in the pre-period before 2012. The row *p-value* refers to the joint test that none of the pre-period interaction effects are significant. Coefficients for post-period interaction terms, economic controls, and fixed effects are not shown. Robust standard errors clustered at the county-level are given in parentheses.

*p $<$ 0.1, **p $<$ 0.05, ***p $<$ 0.01