THROW AWAY THE KEY OR THROW AWAY THE JAIL? The Effect of Punishment on Recidivism and Social Cost

Miguel F. P. de Figueiredo[*]

ABSTRACT

We jail too many people and it costs too much. Incarceration is not only expensive, it also is prone to "hardening" and negative peer learning effects that may increase recidivism. With local, state, and federal budgets at a breaking point, politicians and regulators are increasingly considering alternative approaches to preventing crime. Yet, they face a problem. Studies show that incapacitation is a successful way of reducing crime, yet most scholars and policymakers think that the only way to incapacitate is to incarcerate. This study demonstrates that this assumption is problematic, arguing that we should understand incapacitation along a continuum, with incarceration at one end. This understanding is important because it allows policy makers to think about new ways to avoid the significant social and fiscal costs of jail while at the same time reaping some of the benefits of incapacitation.

This article explores the relationship between incapacitation and incarceration in the context of drunk driving. Policy makers have adopted a variety of incarceration alternatives to curb drunk driving, and this creates a kind of natural experiment that allows for the rigorous testing of the effects of sanctions on future behavior and that derives policy implications for regulating crimes of addiction. This article is the first to examine the

effectiveness of the sanctions in curbing recidivism and vehicle crashes with some 200,000 alcohol tests.

Four key results emerge from the study. First, it demonstrates that non-carceral sanctions can be effective. Second, the primary channel through which drunk driving sanctions are effective in reducing recidivism and crashes is incapacitation, rather than specific deterrence. Third, non-carceral sanctions have varied success based on what form they take and who they target. A law passed mandating victim panels, increasing the length of license suspensions, and stimulating the use of ignition interlock devices (IIDs)—which require the driver of a vehicle to take an alcohol test—reduced crashes during and after suspension of a driver's license. The same law decreased recidivism during the suspension period, but these recidivism-reducing effects ended soon after the license suspension did. In addition, a license suspension enhancement targeting those with higher blood alcohol content levels neither reduced recidivism nor crashes. Fourth, the probability of recidivism and subsequent crashes for first-time offenders given at least 6 to 24 hours of jail, fines, and a license suspension had no statistically significant effect relative to those who had no sanctions. This suggests that drunk driving sanctions at the legal limit are ineffective. This article explains these results, discusses theoretical and legal reform implications, and also outlines a trajectory for improving causal inference in the study of criminal law. This article concludes by discussing the promise and limitations of generalizing from the results to other domains of crime and law.

INTRODUCTION

For the first time in nearly four decades, there is a slow retrenchment of policies that created the prison boom. Although the prison population more than tripled from 220 per 100,000 in the population in 1980 to 732 in 2010,[1] the number of inmates in custody has declined every year since 2010.[2] The U.S. Supreme Court laid some of the groundwork for this decline through several rulings that gave judges greater autonomy with respect to sentencing, and ordering the release of inmates in response to prison overcrowding.[3] Most notably, in *Brown v. Plata*,[4] the Court compelled prison authorities in California to reduce its prison population by some 40,000 inmates in two years, as a result of what the majority viewed as inhumane conditions from prison overcrowding.

Recent public opinion shifts, along with changes at the state and local level, have also contributed to a slow backlash against mass incarceration. Prison spending now exceeds expenditures on education and public health,

---

1. LAUREN E. GLAZE, BUREAU OF JUSTICE STATISTICS, CORRECTIONAL POPULATIONS IN THE UNITED STATES, 2010 8 app. tbl.3 (2011); JOHN SCHMITT ET AL., CTR. FOR ECON. & POLICY RESEARCH, THE HIGH BUDGETARY COST OF INCARCERATION 13 app. tbl.1 (2010). The data for 1980 relies on Bureau of Justice Statistics prisoner and jail inmate custody totals, while population data is taken from the 1980 U.S. Census. The 2010 data includes the total number in custody held in state or federal prisons, or in local jails, as of December 31, per 100,000 U.S. residents as of January 1 of the following year.

2. E. ANN CARSON & DANIELA GOLINELLI, BUREAU OF JUSTICE STATISTICS, PRISONERS IN 2012—ADVANCE COUNTS 1 (2013); *see also* Erica Goode, *U.S. Prison Populations Decline, Reflecting New Approach to Crime*, N.Y. TIMES, July 26, 2013, at A11, http://www.nytimes.com/2013/07/26/us/us-prison-populations-decline-reflecting-new-approach-to-crime.html; Press Release, Bureau of Justice Statistics, U.S. Prison Population Declined for Third Consecutive Year During 2012 (July 25, 2013), http://www.bjs.gov/content/pub/press/p12acpr.cfm.

3. In 2005, the U.S. Supreme Court gave judges greater discretion to depart from mandatory minimums in *U.S. v. Booker*, 543 U.S. 220, 245 (2005), a landmark case that made the Federal Sentence Guidelines advisory. The Court ruled that the Sixth Amendment trial by jury guarantee was inconsistent with the Federal Sentencing Guidelines being mandatory. *Id.* Prior to *Booker*, the Federal Sentencing Guidelines were mandatory for all federal judges to follow, but in *Apprendi v. New Jersey*, 530 U.S. 466, 518, 523 (2000) and *Blakely v. Washington*, 542 U.S. 296, 308 (2004), the Court also limited the force of the mandatory sentencing guidelines based on Sixth Amendment issues. Two years later, in *Kimbrough v. United States*, 552 U.S. 85, 107–11 (2007), the Court determined that there was no rational basis for the Federal Sentencing Guidelines to have a 100-to-1 sentencing disparity ratio when an individual possesses, sells, or traffics crack versus cocaine. Two other decisions worth noting are *Gall v. United States*, 552 U.S. 38, 59 (2007), and *Rita v. United States*, 551 U.S. 338, 357–58 (2007), both of which reaffirmed the importance of the sentencing guidelines in playing a role in sentencing decisions. *See generally* Jelani Jefferson Exum, *The More Things Change: A Psychological Case Against Allowing the Federal Sentencing Guidelines to Stay the Same in Light of Gall, Kimbrough, and New Understandings of Reasonableness Review*, 58 CATH. U. L. REV. 115 (2008).

4. 131 S.Ct. 1910, 1944–47 (2011).

concerning politicians.[5] State budget cuts (especially after the 2008 financial crisis), public opinion shifts amid lower crime rates,[6] the growth of alternative courts,[7] and legislative reforms of sentencing[8] have come together to create a moment where reform might be possible. In light of these events, policymakers more than ever before are seeking alternatives in the wake of a historically polarized policy debate that often pits those emphasizing the need for greater public safety against those who stress fiscal and humanitarian concerns.[9] How can the state control crime effectively without further restraining budgets and worsening prison conditions?

---

    5.   *See, e.g.*, Allegra M. McLeod, *Confronting Criminal Law's Violence: The Possibilities of Unfinished Alternatives,* 8 UNBOUND: HARV. J. LEGAL LEFT 109, 111–12 (2012–2013) (citing SCHMITT ET AL., *supra* note 1) ("[S]pending on prisons outpaces investments in education, early childhood programs, and public health; and family ties are routinely ruptured by criminal law's intervention in ways that contribute to inter-generational cycles of poverty, underemployment, and disadvantage").

    6.   *See, e.g.*, THE MELLMAN GRP. & PUB. OP. STRATEGIES, PUBLIC OPINION ON SENTENCING AND CORRECTIONS POLICY IN AMERICA 1 (2012), http://www.pewtrusts.org/~/media/assets/2012/03/30/pew_nationalsurveyresearchpaper_final.pdf ("1. American voters believe too many people are in prison and the nation spends too much on imprisonment. 2. Voters overwhelmingly support a variety of policy changes that shift non-violent offenders from prison to more effective, less expensive alternatives. 3. Support for sentencing and corrections reforms (including reduced prison terms) is strong across political parties, regions, age, gender, and racial/ethnic groups."); Robert Weisberg & Joan Petersilia, *The Dangers of Pyrrhic Victories Against Mass Incarceration*, DÆDALUS, Summer 2010, at 124, 125, 126 (stating that the public "has exhibited some softening of attitude toward those perceived as nonviolent drug offenders," and that "[p]olling suggests that the public is at least slightly less passionately in favor of prison and long sentences as the solution to the crime problem, especially because we now have less of a crime problem.").

    7.   *See, e.g.*, Allegra M. McLeod, *Decarceration Courts: Possibilities and Perils of a Shifting Criminal Law*, 100 GEO. L.J. 1587, 1649 (2012).

    8.   *See, e.g.*, JUDITH GREENE & MARC MAUER, THE SENTENCING PROJECT, DOWNSCALING PRISONS: LESSONS FROM FOUR STATES 3 (2010); David Cole, *Turning the Corner on Mass Incarceration?*, 9 OHIO ST. J. CRIM. L. 27, 29–31 (2011); Goode, *supra* note 2; Louis Michael Seidman, *Hyper-Incarceration and Strategies of Disruption: Is There a Way Out?*, 9 OHIO ST. J. CRIM. L. 109, 129 (2011); Weisberg & Petersilia, *supra* note 6, at 124.

    9.   *See* Philip J. Cook & Jens Ludwig, *More Prisoners Versus More Crime is the Wrong Question*, 185 BROOKINGS POLICY BRIEF 1 (2011) ("[F]raming the incarceration debate as a tradeoff between public safety and public finance is far too narrow. . . . Crime could actually be reduced if the savings were put to use in strengthening other criminal justice programs and implementing other reforms."); Goode, *supra* note 2 (quoting Adam Gelb, director of the Pew Charitable Trusts' Public Safety Performance Project: "Policy makers are not holding their noses and saying we have to scale back prisons to save money. The states that are showing drops are states that are thinking about how they can apply research-based alternatives that work better and cost less."). The *Brown* decision is characteristic of this polarized debate. *Brown*, 131 S.Ct. at 1923–25. The majority opinion, written by Justice Kennedy, highlighted unconstitutional conditions, including prison overcrowding and inadequate medical care. *Id.* at 1928. In the dissent, Justice Alito stated that the Court was gambling with the safety of the people of California. *Id.* at 1961. In a recent decision on an order in the case, Justice Scalia called the order

Non-carceral sanctions[10] have the potential to reduce prison overcrowding, decrease recidivism, and lower corrections costs. Despite these advantages, some states have been slow to adopt them, in part because their effects are not well documented or understood. In addition, politicians and interest groups advocating "tough on crime" policies favoring increased punishment severity have curried favor with the electorate. Incarceration alternatives also often lack a strong "expressive function"[11]—a clear expression of societal disgust and condemnation—relative to incarceration, creating additional barriers for their adoption. As a result, rigorous evaluation of the effectiveness of non-carceral sanctions has largely been overlooked in the extant literature.[12] Consequently, use of non-carceral sanctions by policymakers and judges has frequently been ad hoc, ex-post, and non-evidence-based, often resulting in ineffective targeting of the sanctions among groups of offenders, and across crimes.[13]

---

a "terrible injunction" that would undermine public safety. Brown v. Plata, 134 S. Ct. 1, 2 (2013) (Scalia, J., dissenting from an order denying a request for a stay). The public safety issue is also raised in *Coleman v. Brown*, 952 F. Supp. 2d 901, 907–09 (E.D. Cal. 2013), part of a number of cases against the State of California for prison conditions that went unaddressed until *Brown*. In *Coleman*, the Eastern District of California, in considering increasing early release of prisoners, examined how the prison population could be reduced "without a significant adverse effect on public safety or the criminal justice system's operation." *Id.* at 907.

10.   These sanctions are also called intermediate sanctions, incarceration alternatives, and alternative sanctions. They include, but are not limited to, community service, electronic monitoring, intensive supervision, sex offender registration, and other protective orders, curfews, fines, "boot camps," and license suspensions.

11.   *See* Dan M. Kahan, *Punishment Incommensurability*, 1 BUFF. CRIM. L. REV. 691, 695 (1998); Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV 591, 601 (1996).

12.   *See, e.g.*, Joan Petersilia et al., *Introduction: The Emergence of Intermediate Sanctions*, *in* SMART SENTENCING: THE EMERGENCE OF INTERMEDIATE SANCTIONS ix, x (James M. Byrne et al. eds., 1992) ("Program descriptions and evaluations [of intermediate sanctions] are scarce and not well publicized."); *Incarceration*, UNIV. CHI. CRIME LAB, http://crimelab.uchicago.edu/page/incarceration (last visited Dec. 18, 2015) ("Intermediate sanctions, such as fines or community service requirements, provide another alternative to incarceration, although the evaluation evidence in this area is relatively limited.").

13.   Although the U.S. Supreme Court has largely praised the U.S. Sentencing Commission for its data-driven approach to sentencing, in *Kimbrough v. United States*, 552 U.S. 85, 96 (2007), the majority used the crack-powder disparity as an example where the Commission "did not use [an] empirical approach in developing the Guidelines sentences." Some rigorous program evaluation efforts of non-carceral sanctions do exist, from which reliable causal inferences about the effects of the sanctions can be made. *See, e.g.*, Joan Petersilia et al., *Intensive Supervision Programs for Drug Offenders*, *in* SMART SENTENCING: THE EMERGENCE OF INTERMEDIATE SANCTIONS, *supra* note 12, at 18, 20; J.J. Prescott & Jonah E. Rockoff, *Do Sex Offender Registration and Notification Laws Affect Criminal Behavior?*, 54 J.L. & ECON. 161, 163–65 (2011). However, the larger trend has been neglect of the rigorous evaluation of incarceration alternatives. *See, e.g.*, Joan Petersilia, *A Decade of Experimenting with Intermediate Sanctions: What Have We Learned?*, *in* 2 NAT'L INST. OF JUSTICE, PERSPECTIVES ON CRIME AND JUSTICE 79, 84–85 (1998) (discussing the problems with previous studies and low participation in alternative

This article begins to fill this important gap. The study reported here is the first to take advantage of a quasi-experiment in Arkansas with the case of drunk driving, an offense that imposes high costs on society, and where the state relies on a variety of non-carceral sanctions. Over the last decade, drunk driving has resulted in 10,000 to 20,000 traffic fatalities nationally per year.[14] One in three people in the United States knows someone who has been in an accident involving a drunk driver, and nearly forty percent of all traffic deaths are related to drunk driving. A study in 2002 estimated the cost to taxpayers and federal and local governments at $51 billion per year, excluding deaths and injuries.[15] The cost is undoubtedly higher today.

To examine whether sanctions are mitigating the costs imposed by drunk driving, this study relies on a research design that reliably estimates the causal effects of sanctions on recidivism and subsequent vehicle accidents. Specifically, this study draws on nearly 200,000 alcohol tests administered after state and local police traffic stops from 2001 until 2013. Drunk driving presents an ideal setting to study the effects of sanctions because of three things that happen simultaneously. First, when individuals drink, they cannot dictate or manipulate their blood alcohol content (BAC) level. Second, there is an artificial but consistent and non-manipulable limit set that determines the adjudication of an offender versus a non-offender; either someone is over the legal BAC limit, or he is not. Third, estimating a driver's BAC is done in a reliable, consistent way without the fear of human manipulation or system-gaming from either the driver or the police.

The goal is to study differences between drivers who are just below the legal BAC level and drivers just at or above it. More precisely, in the tables throughout the paper, I compare first-time offenders who have a BAC of 0.065–0.079 with those with a BAC of 0.08–0.094. I also show the robustness of the results at other ranges close to the legal limit in bandwidth sensitivity plots. Currently, in Arkansas, a person at or just above the legal limit typically spends time in jail, pays a fine and court costs, incurs a 180-day license suspension, and attends a mandatory victims panel; the person just below the legal limit goes free. I then compare the recidivism and subsequent vehicle

---

sanction programs); Weisberg & Petersilia, *supra* note 6, at 127 (describing how previous efforts for alternative sanctions "sometimes proved futile because investment in the logistics and the research basis for the alternative sanctions was often neglected, as if the moral attraction to alternative sanctions caused policy-makers and reformers to ignore the hard and expensive work the sanctions require.").

14.   *See* NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., TRAFFIC SAFETY FACTS: ALCOHOL-IMPAIRED DRIVING 2 fig.1 (2014), http://www-nrd.nhtsa.dot.gov/Pubs/812102.pdf.

15.   LAWRENCE J. BLINCOE ET AL., NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., THE ECONOMIC IMPACT OF MOTOR VEHICLE CRASHES 2000, at 2 (2002), http://www-nrd.nhtsa.dot.gov/Pubs/809446.pdf.

crash rates of these two groups, and also examine the outcomes at a higher BAC level of 0.15, which triggers a longer license suspension. With a large comparison group just below, and just at or above the BAC threshold for the legal limit, one can compare these groups that are statistically indistinguishable with respect to every variable, except for the sanctions, in order to isolate the effect of the "treatment" (the sanctions at the legal limit, in this case) on recidivism and subsequent vehicle accidents. Since it is theoretically likely (and statistically testable with regard to observable pre-existing characteristics of the two groups) that BAC levels in these two different subgroups are non-strategically chosen, the research design emulates a randomized experiment—the "gold standard" for causal inference—close to the BAC threshold that triggers the sanctions.

Three key results emerge from the study. First, the primary channel through which drunk driving sanctions are effective in reducing recidivism and crashes is incapacitation, rather than specific deterrence. Second, non-carceral sanctions have varied success based on what form they take and whom they target. A law passed mandating victims panels, increasing the length of license suspensions, and stimulating the use of ignition interlock devices (IIDs)—which require that the driver of a vehicle take an alcohol test to start or continue driving a vehicle—was effective in reducing crashes both while the driver's license was suspended and after. The same law was effective in reducing recidivism during the suspension period, but recidivism-reducing effects go away afterwards. In addition, a license suspension enhancement targeting those at higher blood alcohol content levels was ineffective in reducing recidivism and crashes. Third, the probability of recidivism and subsequent crashes for first-time offenders given at least six to twenty-four hours of jail, fines, and a license suspension was not statistically distinguishable from the probability of those who received an alcohol test, but no sanctions.

The study addresses three core challenges in the existing literature. First, close analysis of non-carceral sanctions invites the opportunity to re-examine how the purposes of punishment are conceptualized and measured. While a large group of scholars characterize incapacitation as only incarceration, I argue that incapacitation should be considered along a continuum, resulting in a more careful consideration of what kinds of criminal behavior can be prevented through various forms of incapacitation.

Second, the article demonstrates the effectiveness of non-carceral sanctions and explores the challenges of isolating the effects of incapacitation from specific deterrence. Separating out whether crime prevention results from incapacitation or deterrence brings a set of challenges for institutional designers and enforcers. Incapacitation restrains the individual's movement

in some form to prevent recidivism. Deterrence, in turn, measures the extent to which perceptions of the levels and probabilities of future punishment influence an individual's decision not to reoffend.[16] Distinguishing the operative mechanism has important implications for how criminal justice systems handle crime, for our understanding of how individuals are initially induced to or prevented from committing crime, and how convicted offenders respond to sanctions. If incapacitation is found to be the primary channel for achieving crime reduction, all else being equal, that finding could justify increasing spending on prisons and stronger non-carceral forms of incapacitation as a primary means of crime prevention. Alternatively, if specific deterrence is the primary mechanism, a greater focus on punishment severity and informing the public about penalties might be more viable policies to act upon. These two channels are not all-inclusive, nor are they mutually exclusive. However, separating out their effects rigorously is an important step in informing criminal justice policy regarding the types, levels, and targeting of criminal sanctions.

Third, I detail the challenge of making rigorous causal inferences about the effect of sanctions on future behavior. Isolating the causal effect of punishment on behavior is challenging, primarily because of the potential confounding effect of other variables on recidivism and subsequent crashes. In addition, knowing the counterfactual—what would have happened either in the absence of the sanctions (i.e. if an enhanced sentence had not been given to an offender)—is particularly difficult to discern in studies without a comparison group.

This article proceeds by going into greater depth about the importance of studying non-carceral sanctions in Section I. Section II discusses three important challenges for the criminal law and policy field: (1) conceptualizing incapacitation; (2) measuring specific deterrence versus incapacitation; and (3) making causal inferences about the effect of sanctions. The remainder of the paper is dedicated to conducting a quasi-experimental evaluation of the effect of drunk driving sanctions on recidivism and vehicle crashes. Section III provides context on the issue of drunk driving. Section IV discusses the research design of the study. Section V discusses the case selection, providing detail on the legal and enforcement regime in Arkansas, the source of my data. Sections VI and VII describe the data and results, respectively. Section VIII discusses policy and legal reform implications, and Section IX concludes the paper.

---

16. I also accept that other purposes of punishment, including rehabilitation and retribution, are also motivations that guide criminal justice policy that should not be overlooked. This paper primarily examines incapacitation and deterrence as an important input into an analysis that would guide legal and policy reform.

I.   THE IMPORTANCE OF NON-CARCERAL SANCTIONS

Non-carceral sanctions have features that are attractive for a number of reasons. First, incarceration alternatives offer an important policy option to achieve the goals of preventing crime and integrating offenders into society. The effects of this prevention and integration can take place without some of the negative effects of imprisonment. In particular, hardening and criminogenic peer learning in prisons can sow the seeds for recidivism and adverse socioeconomic outcomes. These adverse outcomes include low educational attainment, high unemployment, and ruptured family and community structures for those who are incarcerated. Second, non-carceral sanctions are often less expensive in comparison to incarceration. Incarceration costs to taxpayers and governments are high. The Vera Institute estimated taxpayers spend approximately $31,286 annually per inmate,[17] and a report by the California legislature estimated the annual cost of incarcerating one inmate in 2008 at some $47,000.[18] New York City's Independent Budget Office released a study stating that in 2012, the City paid a staggering $167,731 annually to feed, house, and guard each inmate.[19] Third, incarceration alternatives for some offenses might offer a politically feasible policy route through which the punishment purposes of incapacitation, deterrence, retribution, and rehabilitation are met, while also not subjecting offenders to what some have perceived to be inhumane prison conditions. Fourth, the sanctions give scholars and policymakers the chance to evaluate the effectiveness of new sanctions technologies, and to offer new causal mechanisms that link varied punishment types and mixes of sanctions to differences in recidivism and societal outcomes. Finally, examination of non-carceral sanctions, as I discuss in Section II, leads to reconsideration of how we define and measure incapacitation, and also consider the mechanisms through which these types of sanctions are effective in reducing recidivism and other negative outcomes.

The existing literature focuses overwhelmingly on the effects of incarceration, neglecting the useful role that non-carceral sanctions can play in the mix of sanctions used to control crime. Almost without exception,

---

17. CHRISTIAN HENRICHSON & RUTH DELANEY, VERA INST. OF JUSTICE, THE PRICE OF PRISONS: WHAT INCARCERATION COSTS TAXPAYERS 9 (2012), http://www.vera.org/sites/default/files/resources/downloads/Price_of_Prisons_updated_version_072512.pdf.

18. *How Much Does It Cost to Incarcerate An Inmate?*, LEGISLATIVE ANALYST'S OFFICE, http://www.lao.ca.gov/PolicyAreas/CJ/6_cj_inmatecost (last visited Nov. 19, 2015).

19. Marc Santora, *City's Annual Cost Per Inmate is $168,000, Study Finds*, N.Y. TIMES, Apr. 24, 2013, at A16, http://www.nytimes.com/2013/08/24/nyregion/citys-annual-cost-per-inmate-is-nearly-168000-study-says.html.

previous program evaluations of non-carceral sanctions rely on descriptive statistical work. While descriptive work can be helpful in discerning underlying patterns and correlations, such designs are subject to omitted variable bias and model specification, and often lack a comparison group for estimating the effect sanctions have on recidivism and other outcomes. Research designs with strong causal inference strategies have been extremely rare. These rigorous research designs can shed light on the effectiveness of these varied sanctions in a variety of domains of punitive, retributive, and rehabilitative measures used by the state.

Traditionally, policymakers, judges, and interest group advocates have made non-data-driven judgments about the use of these sanctions. Judges and policymakers, in particular, often make decisions without reliable data on and analysis of the effects of non-carceral sanctions. The need for reliable studies in this area has particularly increased in the post-*Booker* era, in which the U.S. Supreme Court made the federal sentencing guidelines advisory.[20] Although policymakers have discussed the benefit of such sanctions from a cost perspective, one additional advantage is their tendency to facilitate societal integration. As a result, as discussed previously, offenders might not experience hardening and negative peer learning that scholars have mentioned as important drivers of recidivism.

II.   THREE CHALLENGES IN THE LITERATURE

A. *Conceptualizing Incapacitation*

In the United States, incapacitation became the predominant logic for the prison boom, as the rehabilitative model that dominated throughout the 1960s and early 1970s started to wane in popularity.[21] Incapacitation features

---

20. United States v. Booker, 543 U.S. 220, 245 (2005).

21. FRANKLIN E. ZIMRING & GORDON HAWKINS, INCAPACITATION: PENAL CONFINEMENT AND THE RESTRAINT OF CRIME 8–12 (1997); Malcolm M. Feeley & Jonathan Simon, *The New Penology: Notes on the Emerging Strategy of Corrections and its Implications*, 30 CRIMINOLOGY 449, 450 (1992) (describing the "new penology" as involving "the language of probability and risk increasingly [replacing] earlier discourses of diagnosis and retributive judgment"). While Zimring & Hawkins offer a compelling account of a shift in the logic of the prison boom, it is important to acknowledge that scholars state that incapacitation, in combination with deterrence, drove the prison boom. *See, e.g.*, BRUCE WESTERN, THE POLITICS AND ECONOMICS OF THE PRISON BOOM 37–38 (2005) (stating that "Republican governors rejected rehabilitation, expanded prison capacity, and turned the penal system to the twin tasks of incapacitation and deterrence"); Jonathan Simon, *Mass Incarceration: From Social Policy to Social Problem*, *in* THE OXFORD HANDBOOK OF SENTENCING AND CORRECTIONS 28 (Joan Petersilia & Kevin R. Reitz eds., 2012) (describing the abandonment of control in California in the 1970s, and the increase in

prominently as one of the primary purposes of punishment; consequently, much of the policy discourse on crime reduction overwhelmingly focuses on the effectiveness of incarceration. To the best of my knowledge, this study is the first to examine the incapacitation effects of non-carceral sanctions.

In defining incapacitation, a number of scholars have assumed incarceration as a necessary component of incapacitation. Four examples provide *prima facie* evidence in support of this point. Thomas Miles and Jens Ludwig, define incapacitation as "the inability of an incarcerated person to commit additional offenses,"[22] thus, making incarceration a necessary feature of incapacitation. William Spelman similarly incorporates incarceration into his conceptualization of incapacitation: "putting criminals behind bars, where they can not get at the rest of us."[23] Arjan Blokland and Daniel Nagin characterize incapacitation as "the crimes averted by their physical isolation during the period of incarceration,"[24] and David Lee and Justin McCrary also define the term as "the mechanical reduction in crime that occurs when offenders are incarcerated and unavailable to commit additional crimes."[25] All four definitions of incapacitation equate it with incarceration, and discuss the prevention of all crimes against society through the isolation incarceration imposes on an individual.

In this section, I argue that incapacitation can refer to any condition that limits or restrains the movement of a defendant where the state is acting to prevent the individual from reoffending in the present or future.[26] Thus,

incarceration that followed, stating the change "abandoned the focus on rehabilitation in favor of punitive segregation intended to achieve deterrence and, more reliably, incapacitative effects").

22.   Thomas J. Miles & Jens Ludwig, *The Silence of the Lambdas: Deterring Incapacitation Research*, 23 J. QUANTITATIVE CRIMINOLOGY 287, 290 (2007).

23.   WILLIAM SPELMAN, CRIMINAL INCAPACITATION 1 (1994).

24.   Arjan A.J. Blokland & Daniel S. Nagin, *Estimating the Effects of Imprisonment: Intended and Unintended Consequences of Incarceration*, in INCAPACITATION: TRENDS AND NEW PERSPECTIVES 221, 221 (Marijke Malsch & Marius Duker eds., 2012). Other authors describe incapacitation or closely related concepts by incorporating incarceration. *See, e.g.*, PETER W. GREENWOOD & ALLAN ABRAHAMSE, SELECTIVE INCAPACITATION x (1982) (describing ". . . incapacitation effect [as] . . . those crimes prevented while offenders are incarcerated").

25.   David S. Lee & Justin McCrary, The Deterrence Effect of Prison: Dynamic Theory and Evidence 3 (July 2009) (unpublished working paper) http://emlab.berkeley.edu/~jmccrary/lee_and_mccrary2009.pdf.

26.   Notably, Marijke Malsch and Marius Duker argue for a similar more expansive definition of incapacitation, where the term "relates to all sanctions and interventions that aim to impede, restrict or make impossible certain actions, without necessarily being accompanied by measures that aim at other goals and effects, such as retribution, rehabilitation, restoration, et cetera." INCAPACITATION: TRENDS AND NEW PERSPECTIVES, *supra* note 24, at 1, 2. For additional broader uses of incapacitation, see ZIMRING & HAWKINS, *supra* note 21; Jonathan Simon, *Total Incapacitation: The Penal Imaginary and the Rise of an Extreme Penal Rationale in California in the 1970s*, in INCAPACITATION: TRENDS AND NEW PERSPECTIVES, *supra* note 24, at 18–19 (discussing how European governments use incapacitation when referring to probation with

incapacitation, as the term is conceived here, necessarily involves the *degree* to which an individual is monitored and/or restrained, ranging from being completely unable to reoffend, to having limited degrees in their ability to recidivate. Some scholars have characterized "total incapacitation" as "the idea that imprisonment for as long as possible is appropriate whenever an offender poses any degree of risk to the community."[27] In reality, even more extreme measures of incapacitation are possible, such as capital punishment, which truly involves complete and "total" incapacitation of an individual. Thus, even incarceration is a limited form of incapacitation, since crimes can be committed both within and outside detention facilities while an individual is incarcerated. Examples include drug lords and gang leaders running organized crime rings from prisons,[28] inmates committing financial crimes while being detained,[29] and prisoners engaging in phone scams against unsuspecting individuals in society.[30] Moreover, the possibility of escape also limits incarceration from being a form of "total" incapacitation.

My focus, though, is on incapacitation alternatives that are less restrictive than incarceration. I include electronic monitoring, probation, sex offender registration, protective orders, curfews, ignition interlock devices, and license suspension in this definition, since they are all designed to restrict the movement of offenders, and prevent the individual from reoffending.[31] Incapacitation, thus, includes any punishment that limits a person's movement or imposes restraint on an individual's action, including the denial of driving privileges through license suspensions, electronic monitoring of an individual's movements through surveillance measures, and commitment of an individual to a mental health facility.[32]

---

community work orders and electronic monitoring, and also using incapacitation to describe the restrained movement of individuals committed to mental health facilities).

27.   Simon, *supra* note 26, at 18.

28.   *See, e.g.*, David Skarbek, *Governance and Prison Gangs*, 105 AM. POL. SCI. REV. 702, 714 (2011).

29.   *See, e.g.*, Press Release, U.S. Att'y's Office, Federal Prisoner Receives an Additional 184 Months' Imprisonment for Crimes Committed While Incarcerated (Aug. 8, 2013), http://www.fbi.gov/detroit/press-releases/2013/federal-prisoner-receives-an-additional-184-months-imprisonment-for-crimes-committed-while-incarcerated.

30.   John Matarese, *Prision Inmate Scam Targeting Cellphones*, 10 NEWS (Oct. 25, 2012, 7:36   AM),   http://www.10news.com/money/dont-waste-your-money/prison-inmate-scam-targeting-cellphones_28104232.

31.   It is worth noting that others have characterized these forms of punishment as incapacitation. *See, e.g.*, Wim Huisman, *The Application of Administrative Law Against Organized Crime: Refusing and Revoking Licenses as Incapacitation*, *in* INCAPACITATION: TRENDS AND NEW PERSPECTIVES, *supra* note 24, at 185, 185.

32.   For an extensive treatment of probation supervision, home confinement, and electronic monitoring, see generally SMART SENTENCING: THE EMERGENCE OF INTERMEDIATE SANCTIONS, *supra* note 12.

Embedded in nearly all of the definitions of incapacitation, either explicitly or implicitly, is the notion of a counterfactual. Specifically, the person being incapacitated is prevented from committing crimes that could have been committed had he or she not been subject to the incapacitation. Yet, the precise nature of the counterfactual is difficult to specify because it depends in large measure on the purpose of the incapacitation—whether, for instance, the individual is being restrained from action to prevent any kind of crime, whether he or she is being isolated from society, or whether the individual is being restrained exclusively from the crime he or she originally committed. Alana Barton describes the difficulty of precisely specifying this counterfactual, which she describes as "hypothetical crimes."

Unlike with retribution, inherent in the theory of incapacitation is a notion of societal risk, punishment is not concerned with the nature of the offender, as is the case with rehabilitation, or with the nature of the offense, as is the case with retribution. Rather, punishment is justified by the risk individuals are believed to pose to society in the future. As a result, individuals can be punished for "hypothetical" crimes. In other words, they can be incarcerated, not for crimes they have actually committed but for crimes it is anticipated or assumed they will commit.[33]

Not only are these "hypothetical crimes" a key part of what drives the punishment decision, they also shape a society's notion of risk,[34] along with media-fueled insecurity about public safety and the salience of "tough on crime" rhetoric in the political arena. Figure 1 shows a conceptualization of incapacitation along a continuum, with varied purposes for incapacitation, the forms of punishment that incapacitation can take, and the category of the punishment mechanism. In discussing purpose, it is important to note that all punishments for incapacitation can also have retributive, deterrent, rehabilitative, and other purposes. The four purposes I highlight—preventing the individual from committing any crime, isolating the individual from society, monitoring or restraining the individual's movement within society, and preventing the individual from committing the individual crime he or she committed previously—are meant to be non-mutually-exclusive "ideal types."[35] As a result, I placed a number of punishments that involve

---

33. Alana Barton, *Incapacitation Theory*, *in* ENCYCLOPEDIA OF PRISONS & CORRECTIONAL FACILITIES 463 (Mary Bosworth ed., 2005).

34. *See, e.g.*, Loïc Wacquant, *Deadly Symbiosis: When Ghetto and Prison Meet and Mesh*, *in* MASS IMPRISONMENT: SOCIAL CAUSES AND CONSEQUENCES 82, 82 (David Garland ed., 2001) (describing how adverse targeting resulting from risk perceptions creates extremely high levels of "hyper-incarceration" in certain communities); *see also* Seidman, *supra* note 8, at 129.

35. Worth noting is an important line of research that offered the concept of "selective incapacitation," which, unlike the more traditional form of incapacitation, is characterized by assigning punishment strictly based on the risk the individual poses to society by reoffending.
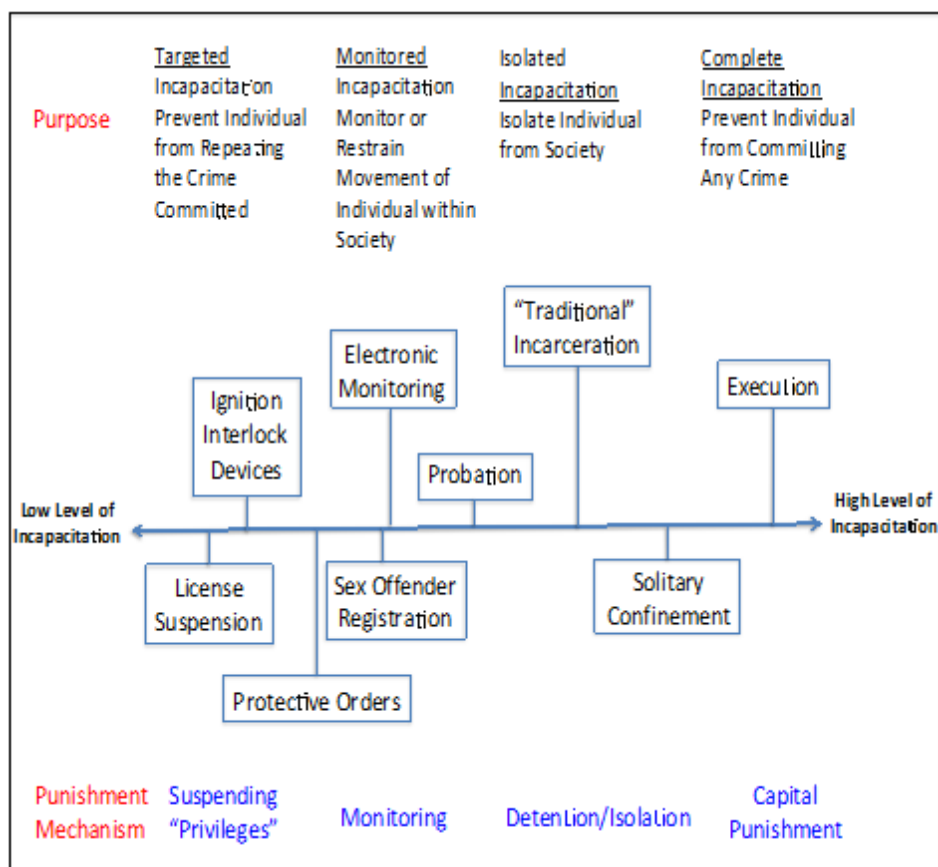
monitoring, such as probation and electronic monitoring between the two purposes.

I suggest the use of four subtypes for incapacitation to help clarify these four purposes: targeted incapacitation, monitored incapacitation, isolated incapacitation, and complete incapacitation. Targeted incapacitation refers to preventing the individual from recommitting the crime that led to the punishment. In the case of drunk driving, ignition interlock devices (IIDs) are a clear example of this type of incapacitation. IIDs are typically installed on the dashboard or steering wheel of the vehicle, and require that a person take a breath test under a pre-specified limit in order to start the car. The Borg Warner Company introduced the first IID in 1970, but they did not become popular until the 1990s when several new features became standard and the National Highway Traffic Safety Administration (NHTSA) developed standards for the devices in 1992.[36] In the past, their ability to incapacitate previous offenders from drunk driving was limited since testing with the device was only required to start the vehicle; a second person could blow into the IID, and allow the offender to drive. Today's devices are more sophisticated, and there are IIDs that require breath tests while the individual is driving, and randomize the timing of when the test is required. Figure 5(a) in Appendix I(A) shows an IID attached to the dashboard of a vehicle, and Figure 5(b) shows a person taking an IID breath test. IIDs are discussed in greater detail in Section IV(A). These devices offer a clear example of targeted incapacitation, since they are designed to prevent a previous offender from committing the crime the person originally committed (in this case, drunk driving). Another sanction in this category is license suspension, which is designed to target the behavior of erratic driving of any form.[37]

Unlike more traditional forms of incapacitation, selective incapacitation does not sentence based on the crime committed, but rather based on the risk profile of the individual, irrespective of the crime.

36. Robert Voas & Paul Marques, *History of Alcohol Vehicle Interlock Programs: Lost Opportunities and New Possibilities*, PACIFIC INSTITUTE FOR RESEARCH AND EVALUATION (2014), http://www.researchgate.net/publication/242204156_History_of_Alcohol_Vehicle_Interlock_Pr ograms_Lost_Opportunities_and_New_Possibilities.

37. Although not commonly practiced in the United States at present, involuntary sterilization (also referred to as forced sterilization or compulsory sterilization) and chemical castration present additional examples of targeted incapacitation. However, the practice, when used against those convicted of rape, may still not incapacitate a person from committing that crime. Involuntary sterilization also reportedly took place targeting the mentally handicapped, mentally ill, the hearing and visually impaired, and epileptics, as part of a eugenics movement that was conducted under the laws in a number of different states. These forms of punishment would not qualify as targeted incapacitation. Other U.S. sterilization programs have targeted prisoners and racial minorities, including African Americans and Native Americans. Although the practice continues today among individuals and groups, the only government known to conduct involuntary sterilization today is Uzbekistan. Natalia Antelava, *Uzbekistan's Policy of*

Figure 1: The Incapacitation Continuum

License suspension is also an example of a second category of incapacitation—monitored incapacitation. Although in the United States license suspensions (and revocations) are primarily used for vehicle-related offenses, some states use them for non-vehicle related offenses. Virginia, for example, revokes driving privileges for making a bomb threat, and non-vehicle related drug offenses.[38] Massachusetts issues license suspensions for

Secretly Sterilising Women, BBC News (Apr. 11, 2012), http://www.bbc.co.uk/news/magazine-17612550. For additional information on compulsory sterilization, see generally Daniel J. Kevles, In the Name of Eugenics: Genetics and the Uses of Human Heredity (1985); Caroline M. Wong, *Chemical Castration: Oregon's Innovative Approach to Sex Offender Rehabilitation, or Unconstitutional Punishment?*, 80 Or. L. Rev. 267 (2001); Paul Lombardo, *Eugenic Sterilization Laws*, Image Archive on the Am. Eugenics Movement, http://www.eugenicsarchive.org/html/eugenics/essay8text.html (last visited Jan. 24, 2016).

38. *Reinstating Driving Privleges*, Va. Dep't of Motor Vehicles, http://www.dmv.virginia.gov/drivers/#reinstate.asp (last visited Oct. 30, 2015).

failure to pay child support, non-vehicle related drug offenses, state income tax violations, outstanding arrests, and failure to register as a sex offender.[39] In the case of motor vehicle violations, license suspensions are designed to prevent the offender from engaging in dangerous driving. However, other purposes, including restraining an individual's movement and retribution, are likely to be at work related to offenses like failure to pay child support or taxes. Other sanctions, such as protective orders, sex offender registration, electronic monitoring, and probation also clearly fall under this purpose of incapacitation.

Isolated incapacitation falls in line with the more common conceptualization and operationalization of incapacitation. Isolated incapacitation involves the separation of the individual from society for the purpose of minimizing (but not eliminating) the risk of recidivism for any crime. Sanctions in this category include some form of incarceration. Although more extreme forms of incarceration are possible, such as solitary confinement, the possibility of a person reoffending still make it a limited form of incapacitation.[40]

Only capital punishment is a truly "complete" form of incapacitation in which the individual is unable to reoffend. More extreme forms of punishment such as dismemberment might make recidivism of certain crimes—and possibly any crimes—impossible while the person is under custody. In these cases, these forms of punishment may also qualify as complete incapacitation.

Two caveats merit attention with respect to categories of purposes on the continuum. First is the issue Barton raises about "hypothetical crimes," or the crimes for which the incapacitation is targeted at preventing.[41] Looking at the purposes on the continuum, on one end, we see that specific incapacitation is directed at preventing the specific crime the individual committed, and on the other end total incapacitation is directed at preventing the offender from committing any future crime. These purposes are intended to be extreme endpoints of a continuum, with the location of punishment form between those two endpoints being determined by the number and severity of crimes at which the incapacitation is directed. Second, although scholars such as Jonathan Simon include incarceration in what he terms "total

---

39. COMMONWEALTH OF MASS., DRIVER'S MANUAL 50–51 (2013) http://www.massrmv.com/rmv/dmanual/Drivers_Manual.pdf.

40. Michael Montgomery, *Gangs Reach Out of Prison to Commit Crimes*, NATIONAL PUBLIC RADIO (Mar. 7, 2005, 3:34 PM), http://www.npr.org/templates/story/story.php?storyId=4525733 (describing how gang members in the Pelican Bay State Prison in California's isolated supermax unit run street gangs from the facility).

41. Barton, *supra* note 33, at 463.

incapacitation,"[42] I believe, for the reasons previously stated, that this term should be exclusively used to describe execution or extreme forms of punishment, such as certain forms of dismemberment, that make an individual physically incapable of reoffending.

The level of incapacitation refers to the degree of restraint against an individual's freedom of movement each punishment typically represents. I acknowledge a degree of subjectivity in placing various punishments along this continuum, especially since sub-categories of the punishments and individual circumstances might make certain punishments more restrictive than others in ostensibly harsher punishment categories. For example, bail could vary greatly in the degree to which it imposes restraint on an individual, based on an individual's willingness to pay, access to funds, income level, and the number of days in jail the individual faces as an alternative, among other factors. Despite this issue, I believe the categories can still serve an important purpose in providing a more refined notion of incapacitation.

The conceptualization of incapacitation is critical in informing decisions about the measurement and operationalization of variables in an empirical analysis, and also provides a framework for thinking about the purposes of incapacitation with non-carceral sanctions. A number of implications emerge from the conceptualization. First, a broader conceptualization of incapacitation inclusive of a wider array of non-carceral sanctions is important. Research that investigates the effectiveness of these punishments not only informs public policy, but also reflects risks that policymakers are willing to take with the integration of offenders in society with various punishments and reveals the reasoning behind their punishment. Second, with a more granular notion of incapacitation, we then are able to focus on measuring incapacitation effects in different ways. This would motivate a research agenda about the effectiveness of different forms of monitoring, and its effects on crime, and other outcomes of substantive interest to both scholars and policymakers.

B. *Untangling Incapacitation from Specific Deterrence*

Even with a more refined notion of incapacitation, another important challenge remains: separating out incapacitation from specific deterrence. Specific deterrence, which refers to how the threat of future sanctions stops an individual criminal from reoffending, is an important mechanism that has an effect on regulating crime. If specific deterrence is working, then the focus

---

42. *See* Jonathan Simon, *How Should We Punish Murder*, 94 MARQ. L. REV. 1241, 1254 n.50 (2010).

of policy should be oriented toward increasing sanction levels and punishment probabilities. If incapacitation is the primary means through which crime is being reduced, restraining the individual's movement in some form is what prevents recidivism and reduces crime.

Distinguishing what mechanism is primarily at work has important implications for how criminal justice systems will handle crime, and for how criminal behavior can be altered. Separating out incapacitation versus specific deterrence effects in a rigorous causal manner is an important step in informing policy regarding the types, levels, and targeting of criminal sanctions. The effectiveness of each also has important implications for budget allocations toward crime reduction and maintaining public safety, and also can have an important social impact on the friends and family of the offender, along with future prospects of education, employment, and health.[43]

Isolating these effects is challenging for a number of reasons. First, changes that increase punishment affect both deterrence and incapacitation at the same time.[44] Second, confounding is also introduced by changes in policy and society taking place at the same time as when a sentence enhancement is enacted, making it challenging to determine the causal effect of sentencing.[45] Third, the absence of a counterfactual that would shed light on what would happen if such a policy were not enacted also complicates making the outcome directly attributable to the punishment policy.

This paper is, by no means, the first to attempt to isolate the effects of deterrence and incapacitation in a causal manner. Daniel Kessler and Steven D. Levitt, to the best of my knowledge, were the first to separate out incapacitation from deterrence by using California's Proposition 8, which increased the severity of sentences for repeat offenders for some crimes, but not others.[46] Of particular relevance to this study is Lee and McCrary's *The Deterrence Effect of Prison: Dynamic Theory and Evidence*, which relies on the same quasi-experimental design to measure the effects of harsher penalties in adult versus juvenile sentencing regimes on recidivism in

---

43.   *See* Emily G. Owens, *More Time, Less Crime? Estimating the Incapacitative Effect of Sentence Enhancements*, 52 J.L. & ECON. 551, 552, 568–72 (2009).

44.   Daniel Kessler & Steven D. Levitt, *Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation*, 42 J.L. & ECON. 343, 343–44 (1999).

45.   *See* Owens, *supra* note 43, at 552.

46.   *See* Kessler & Levitt, *supra* note 44, at 345, 348; *see also* David S. Abrams, *Estimating the Deterrent Effect of Incarceration Using Sentencing Enhancements*, 4 AM. ECON. J. 32, 33 (2012) (citing Francesco Drago et al., *The Deterrent Effects of Prison Evidence from a Natural Experiment*, 117 J. POL. ECON. 257 (2009); Eric Helland & Alexander Tabarrok, *Does Three Strikes Deter? A Non-Parametric Investigation*, 42 J. HUM. RESOURCES 309 (2007)); Owens, *supra* note 43, at 552–53

Florida.[47] Unlike Kessler and Levitt, who find a deterrence effect and no incapacitation effect,[48] Lee and McCrary essentially find the opposite—that the main mechanism through which penalties are having an effect is incapacitation, rather than deterrence.[49]

This research examines incapacitation and deterrence effects for non-carceral sanctions. The policy consequences of non-carceral sanctions for both incapacitation and deterrence are significant. If we find that there are incapacitation effects for non-carceral sanctions, the sanctions might achieve similar effects at a fraction of the cost of incarceration, and the sanctions could be used to reduce prison overcrowding. Relatedly, the incarceration alternatives examined might also reduce some of the negative effects of incarceration, including criminogenic peer learning and hardening, which are thought to increase recidivism. If deterrence effects are found, then the introduction of lower threshold penalties could potentially lead to more effective targeting of sanctions. There could be a "ratcheting down" of sanctions for those that lie close to the margin for incarceration, which could result in a more effective use of marginal deterrence.[50] Moreover, the paper not only examines incapacitation effects on recidivism, but it also examines the effects on vehicle crashes, another outcome with important societal consequences.

C. *Moving from Correlation to Causation: Making Causal Inferences About the Effects of Punishment*

Determining the causal effect of crime and corrections policy brings its own set of challenges for legal scholars, policymakers, criminologists, and other social scientists. Although previous studies have examined the determinants of recidivism, most have done so in a correlational or predictive manner using various forms of regression analyses to control for factors such

---

47. *See* Lee & McCrary, *supra* note 25, at 32.

48. *See* Kessler & Levitt, *supra* note 44, at 349.

49. Although Lee and McCrary conclude that "if lengthening prison sentences leads to significant crime reduction, it is likely operating through a direct, 'mechanical' incapacitation effect, rather than through a behavioral response to the threat of punishment," they also state that "deterrence elasticities with respect to sentence lengths are no more negative than -0.13 for young offenders." *See* Lee & McCrary, *supra* note 25, at 1, 4.

50. Marginal deterrence is the idea that the severity of the crime committed or the number of crimes committed should determine the level of punishment, so that offenders who commit more severe and/or numerous crimes should be punished more severely. The idea is that the presence of marginal deterrence would properly incentivize offenders and would-be offenders not to benefit from committing additional crimes, including future crimes of greater severity.

as criminal history, which could lead to recidivism.[51] While these approaches with observational data can show correlation, the results often depend on a number of strong assumptions, especially when making causal inferences about the effect of an intervention on an outcome.

There is a relatively strong consensus among econometricians, statisticians, and other quantitative social scientists that regression results from observational data, without some form of random variation (also referred to as exogenous variation), are highly sensitive.[52] The results often impose a number of modeling assumptions on the data, especially because they can be very sensitive to unobserved factors that can drive the results. Experimental and quasi-experimental methods offer ways of dealing with these important issues.[53]

To facilitate understanding of the importance of causal inference and the pitfalls of regression, a short discussion of the Neyman-Rubin Causal Model[54]

---

51. *See, e.g.*, Ted Chiricos et al., *The Labeling of Convicted Felons and its Consequences for Recidivism*, 45 CRIMINOLOGY 547, 560 (2007); Douglas A. Smith & Patrick R. Gartin, *Specifying Specific Deterrence: The Influence of Arrest on Future Criminal Activity*, 54 AM. SOC. REV. 94, 100–01 (1989); Cassia Spohn & David Holleran, *The Effect of Imprisonment on Recidivism Rates of Felony Offenders: A Focus on Drug Offenders*, 40 CRIMINOLOGY 329, 334–36 (2002); Yan Zhang et al., *The Impact of State Sentencing Policies on the U.S. Prison Population*, 37 J. CRIM. JUST. 190, 191 (2009).

52. *See, e.g.*, Daniel E. Ho & Donald B. Rubin, *Credible Causal Inference for Empirical Legal Studies*, 7 ANN. REV. L. SOC. SCI. 17, 26 (2011) ("For causal inference, the overwhelming recognition in applied statistics is that regression alone is fragile."); *see generally* RICHARD BERK, REGRESSION ANALYSIS: A CONSTRUCTIVE CRITIQUE (2004); CHARLES F. MANSKI, IDENTIFICATION PROBLEMS IN THE SOCIAL SCIENCES (1995); DONALD B. RUBIN, MATCHED SAMPLING FOR CAUSAL EFFECTS (2006); Joshua D. Angrist & Alan B. Krueger, *Empirical Strategies in Labor Economics*, *in* HANDBOOK OF LABOR ECONOMICS 1277 (Orley Ashenfelter & David Card eds., 1999); Rajeev H. Dehejia & Sadek Wahba, *Causal Effects in Non-Experimental Studies: Reevaluating the Evaluation of Training Programs*, 94 J. AM. STAT. ASS'N 1053 (1999); Daniel E. Ho et al., *Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference*, 15 POL. ANALYSIS 199 (2007); Gary King & Langche Zeng, *The Dangers of Extreme Counterfactuals*, 14 POL. ANALYSIS 131 (2006); Robert J. Lalonde, *Evaluating the Econometric Evaluations of Training Programs with Experimental Data*, 76 AM. ECON. REV. 604 (1986); Donald B. Rubin, *Bayesian Inference for Causality: The Importance of Randomization*, 6 ANNALS STAT. 34 (1978) [hereinafter Rubin, *Bayesian Inference*]; Donald B. Rubin, *Matching to Remove Bias in Observational Studies*, 29 BIOMETRICS 159 (1973) [hereinafter Rubin, *Matching to Remove Bias*]; Jeff Strnad, *Should Legal Empiricists Go Bayesian?*, 9 AM. L. ECON. REV. 195 (2007).

53. Experiments are but one of a number of different methods used to make causal inferences about the effect of X on Y. In addition to a host of quasi-experimental methods, including the two used in this paper, there are also qualitative methods and deductive theories one can use to produce causal inferences.

54. The Neyman-Rubin Causal Model, is also called the Neyman-Holland-Rubin Causal Model, the Rubin Causal Model, and the potential outcomes framework for causal inference. I draw heavily on Jasjeet Sekhon's description of the history and technical notation of the Neyman-Holland-Rubin Causal Model. For a more detailed history of the model, see Jasjeet S. Sekhon,

is in order. The model clarifies a precise approach to causation, which we can then examine in the context of sentencing. Jerzy Neyman first developed the idea of a potential outcomes framework in which each observation in the study had two potential outcomes; it could be assigned to either a treatment or a control group.[55] The causal effect is defined by the difference between these two potential outcomes, but we are unable to observe one of these outcomes, since the same unit of analysis cannot travel back in time and experience the counterfactual.[56] In an influential paper, Paul Holland dubbed this the "fundamental problem of causal inference."[57] As a result, we never directly observe and measure a causal effect, but only make *causal inferences* about the effect of some treatment or intervention on an outcome. In a series of papers, William G. Cochran and Donald Rubin later developed a framework for thinking about the Neyman model with application to research with observational data.[58] For an experimental research design to be executed effectively, the design requires at a minimum (1) specification of the treatment and control; (2) random assignment of the treatment to the randomization group; and (3) numerosity of observations in the treatment and control groups.[59] The random assignment of the treatment is what allows one

---

The Neyman-Rubin Model of Causal Inference and Estimation via Matching Methods, *in* OXFORD HANDBOOK OF POLITICAL METHODOLOGY 271, 272–73 (Janet Box-Steffensmeier et al. eds., 2008).

55. *See* Jerzy Splawa-Neyman et al., *On the Application of Probability Theory to Agricultural Experiments*, 5 STAT. SCI. 465, 471–72 (1990).

56. More formally, if we let Y_{it} denote the potential outcome for unit i if that unit receives treatment t, and Y_{ic} denotes the potential outcome for the same unit in the control group c, then the treatment effect, τ_i is defined by $\tau_i = Y_{it} - Y_{ic}$.

57. Paul W. Holland, *Statistics and Causal Inference*, 81 J. AM. STAT. ASS'N. 945, 947–48 (1986).

58. *See generally* RUBIN, *supra* note 52; William G. Cochran, *Matching in Analytical Studies*, 43 AM. J. PUB. HEALTH & NATION'S HEALTH 684 (1953); William G. Cochran, *The Planning of Observational Studies of Human Populations*, 128 J. ROYAL STAT. SOC'Y., SERIES A 234 (1965); Rubin, *Matching to Remove Bias*, *supra* note 52, at 159; Donald B. Rubin, *Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies*, 66 J. EDUC. PSYCHOL. 688 (1974); Donald B. Rubin, *Multivariate Matching Methods that are Equal Percent Bias Reducing, I: Some Examples*, 32 BIOMETRICS 109 (1976); Donald B. Rubin, *Multivariate Matching Methods That are Equal Percent Bias Reducing, II: Maximums on Bias Reduction for Fixed Sample Sizes*, 32 BIOMETRICS 121 (1976); Rubin, *Bayesian Inference*, *supra* note 52; Donald B. Rubin, *Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies*, 74 J. AM. STAT. ASS'N. 318 (1979).

59. Although experiments can be done with relatively small numbers of observations, the point is that ex-ante, most researchers would like to have sufficient statistical power to detect a treatment effect. In order to obtain a treatment effect, the number of observations in the treatment and control groups must be sufficiently large, such that if we let δ_0 denote the expected treatment effect size, α = significance level (or the probability of a Type I error), β = a given power level, s = standard deviation, and z = the $1 - \beta$ quartile of the normal distribution, we can use the following formula to calculate the statistical power needed to obtain a treatment effect: $\delta_0 =$

to make strong causal inferences about the effects of the intervention on the outcome, because, in expectation, all unobserved factors are balanced across the treatment and control groups. I say "in expectation" because in order for all unobservable factors uncorrelated with the treatment to be "controlled for," the randomization has to have "worked." There is an expectation of equivalence across the treatment and control groups because across multiple random draws, there will be equivalence across the groups on all observed and unobserved variables. However, any one given draw may not achieve equivalence. In practice, experimenters typically verify whether there is equivalence on observed variables across the treatment and control groups in the data to see if the randomization at least worked for those factors that can be observed.[60] Often, the most important variable to have equivalence on in experimental work is the lagged outcome (e.g. if one analyzed the effect of an intervention on recidivism, one would search for equivalence across the treatment and control groups based on the individuals' criminal histories).

Experiments do come with their own set of drawbacks. The main one, for our purposes, is the issue of external validity—the ability to generalize from the results across time, context, and alternative realizations of treatments. Typically, there are two approaches to dealing with the issue. The first is using replication of experiments across various times, contexts, and realizations in order to examine the stability of the results. The second is to develop theories of equivalence and expected results that would allow one to generalize beyond the randomization group.

In the case of examining the effect of sanctions at the legal limit or increased punishment, the treatment is the enhanced sentence,[61] relative to a

---

$\left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$. It is worth noting that these conditions are necessary, but not sufficient, for an effective experimental design to be executed. The discussion is circumscribed to these conditions in order to make the points necessary for the substantive discussion.

60. Debate exists in the literature about whether one can and should control for factors after an experiment has been conducted, in the event of not having equivalence (also known as balance) on a variable (in this case, referred to as a covariate) across the treatment and control groups. *See, e.g.*, David A. Freedman, *On Regression Adjustments to Experimental Data*, 40 ADVANCES APPLIED MATHEMATICS 180 (2008); Winston Lin, *Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique*, 7 ANNALS APPLIED STAT. 295 (2013); Donald P. Green, Regression Adjustments to Experimental Data: Do David Freedman's Concerns Apply to Political Science? 20–22 (July 8, 2009) (paper presented at the annual meeting of the Society for Political Methodology).

61. A sentence enhancement increases the severity of a punishment based on some established criterion or "trigger." Sentence enhancements are typically increased either because of a prior conviction or because of the more serious nature of a particular offense. For an in-depth examination of recidivist enhancements, see Sarah French Russell, *Rethinking Recidivist Enhancements: The Role of Prior Drug Convictions in Federal Sentencing*, 43 U.C. DAVIS L. REV. 1135, 1143–46 (2010).

control group that receives a reduced sentence.[62] Under the potential outcomes framework, we would ideally want to observe the same individual at the same time receiving and serving both the enhanced and "normal" sentences. Because of the fundamental problem of causal inference, there is an impossibility of observing the counterfactual outcome (e.g., if the person was assigned to the treatment group and received the sentence enhancement, it is not possible to know what would have happened if the same person had received the reduced sentence, and vice versa). Consequently, as I stated earlier, we can only make causal inferences about the effect of the sentence enhancement (the "treatment") by attempting to simulate the counterfactual and discuss the results in probabilistic terms.[63]

A number of factors specific to evaluating the effects of criminal sanctions make it challenging to arrive at causal inferences. First, the severity of punishments can interact with a number of other factors simultaneously, creating difficulties in isolating the effects of the marginal sentence severity. Second, discretion in the criminal system—especially police profiling; prosecutorial discretion to charge, bargain, or drop cases; and judicial discretion in sentencing—can result in selection effects, complicating the determination of the effects of treatments on outcomes. Third, unobserved characteristics of a defendant that lead to harsher sentences may also have an impact on the defendant's probability of recidivating, and also for our study, getting into a subsequent vehicle crash. Taken together, these and other factors make designing and conducting a study with credible causal inference a rather challenging endeavor to undertake.

III.   THE CASE OF DRUNK DRIVING

Sentencing for drunk driving provides an opportunity to examine a number of these challenges in the literature on crime and criminal law, while also applying the conceptualization of incapacitation previously discussed. First, drunk driving sentencing is applied in a manner that enables strong causal inferences to be made about the effects of punishment. As was discussed previously, studies of the effects of criminal sanctions frequently suffer from the problems of selection bias. Sentences are rarely, if ever, randomly assigned, and so the difficulty arises in making causal inferences

---

62.   Ho and Rubin offer a similarly helpful example from the literature on the effect of prison conditions on recidivism to illustrate the point. Ho & Rubin, *supra* note 52, at 22–26.

63.   Additional assumptions in the model include no stable unit treatment value assumption (SUTVA), which assumes no "contamination" of the units in the treatment group either through, for example, contact with someone in the control group, affecting their outcomes. In addition, there cannot be "hidden" administration of the treatment. *See id.* at 21.

about the effect of the sentence on behavior when there are a wide range of unobservable variables that could be driving recidivism and subsequent crashes. With drunk driving, sentencing in most jurisdictions is principally determined by BAC. The formulaic, non-discretionary, and quantifiable aspects of the sentence make it amenable to a quasi-experimental design. This design allows for a more rigorous study of the effects of punishments, since unobservable factors correlated with the assignment of a set of sanctions (whether they are at the legal limit or are an enhancement) are taken care of (in expectation) with the quasi-experimental feature of the research design.

Second, when sentencing for drunk driving offenses, judges can choose a variety of sanctions along the incapacitation continuum, including incarceration, ignition interlock devices (IIDs),[64] vehicle impoundment, and license suspensions. IIDs require the driver to be breathalyzed before operating a vehicle.[65] If the driver's BAC is above the permitted level set by a court or administrative agency, the device can prevent the vehicle from starting.[66] To prevent non-drivers from giving samples, more recent IIDs require the driver to give breath samples while he or she is driving at randomly determined time intervals (typically between five minutes and one hour). Some IIDs also photograph drivers while a breath sample is given. Despite these safeguards, offenders can still circumvent IIDs by driving a vehicle without a device.[67]

Two other alternative sanctions—license suspensions[68] and vehicle impoundment—rely on weaker and stronger forms of incapacitation, respectively, than IIDs. Like IIDs, license suspensions also attempt to prohibit a DWI offender from driving altogether. A number of empirical studies have found that license suspensions are effective in reducing DWI recidivism.[69] Peck, et al., for instance state "[T]here is no question that license suspensions have a significant effect in reducing the accident and drunk driving frequency of convicted DUI offenders . . . the overall consistency of

---

64. The devices are also referred to as breath alcohol ignition interlock devices (BAIID) or simply ignition interlock devices. *See* NAT'L. HIGHWAY TRAFFIC SAFETY ADMIN., IGNITION INTERLOCKS—WHAT YOU NEED TO KNOW 1 (2d ed. 2014) [hereinafter NHTSA REPORT].

65. *Id.*

66. *Id.*

67. *Id.* at 3.

68. For the sake of simplicity, we group license suspension and revocation under the same umbrella.

69. *See, e.g.*, Robert B. Voas, *Evaluation of Jail as a Penalty for Drunken Driving*, 2 ALCOHOL DRUGS & DRIVING 47 (1986); R. C. Peck et al., *The Comparative Effectiveness of Alcohol Rehabilitation and Licensing Control Actions for Drunk Driving Offenders: A Review of the Literature*, 1 ALCOHOL DRUGS & DRIVING 15 (1985); James L. Nichols & H. Laurence Ross, *The Effectiveness of Legal Sanctions in Dealing with Drinking Drivers*, 6 ALCOHOL DRUGS & DRIVING 33 (1990).

the results from different investigators, using different quasi-experimental designs, precludes any other conclusion."[70]

Yet, despite the extant literature, license suspensions still remain a relatively weak form of incapacitation, since enforcement of the sanction is challenging and largely passive, with stops for other offenses, or at sobriety checkpoints, being the main means through which license suspensions are enforced.[71] To date, very few scholars have examined the causal effect of license suspensions on recidivism. Realizing the limitations of license suspensions, a number of states enacted more aggressive laws targeting the vehicle. Sanctions have included registration cancellation, special license plates or plate stickers for DWI offenders,[72] license plate confiscation,[73] vehicle impoundment, and vehicle forfeiture.[74] These programs, which were mostly targeted at the most egregious offenders, resulted in reducing recidivism,[75] and were also seen as draconian in the costs they imposed on the offender and his or her family.[76] Because they were not viewed as a policy that could be applied broadly, and because of the availability of IIDs as an alternative, other vehicle-based sanctioning has declined in recent years. Taken together, the varying degree of incapacitation, and the variety of non-carceral sanctions used in DWI sentencing offer an important first step in the study of this type of punishment.

Fourth, along with the wide range of incapacitation, DWI is also expansive in its prevalence, not only in the more than one million arrested every year for impaired driving, but also in the range of socioeconomic groups arrested for DWI. Although those arrested are overwhelmingly male (approximately 70–80%), offenders are heterogeneous in terms of race, income, and region, allowing for the study of heterogeneous effects across these subgroups.

---

70. Peck et al., *supra* note 69, at 57.
71. Robert B. Voas et al., *Temporary Vehicle Immobilization: Evaluation of a Program in Ohio*, 29 ACCIDENT ANALYSIS PREVENTION 635, 635 (1997).
72. R. B. VOAS & A. S. TIPPETTS, U.S. DEP'T TRANSP., ASSESSMENT OF IMPOUNDMENT AND FORFEITURE LAWS FOR DRIVERS CONVICTED OF DUI, PHASE II REPORT: EVALUATION OF OREGON AND WASHINGTON VEHICLE PLATE ZEBRA STICKER LAWS 5–6 (1994); H. Laurence Ross et al., *License Plate Confiscation for Persistent Alcohol Impaired Drivers*, 28 ACCIDENT ANALYSIS & PREVENTION 53, 53 (1996).
73. Voas et al., *supra* note 71, at 635–36.
74. For a comprehensive survey of vehicle-targeted DWI sanctions, see ROBERT B. VOAS, U.S. DEP'T TRANSP., ASSESSMENT OF IMPOUNDMENT AND FORFEITURE LAWS FOR DRIVERS CONVICTED OF DWI PHASE I REPORT: REVIEW OF STATE LAWS AND THEIR APPLICATION 48–57 (1992).
75. Randy W. Elder et al., *Effectiveness of Ignition Interlocks for Preventing Alcohol-Impaired Driving and Alcohol-Related Crashes: A Community Guide Systematic Review*, 40 AM. J. PREVENTIVE MED. 362, 362–63 (2011).
76. *Id.* at 363.

Finally, although the study of drunk driving provides a number of advantages that for theoretical and conceptual reasons make it amenable to a rigorous research design, the crime itself has serious practical implications for society. Despite active interest group campaigns and steep penalties, drunk driving remains a relatively widespread phenomenon in the United States. NHTSA reported that 32,885 driving-related fatalities took place in 2010, and 10,228 of them—a staggering 31.1%—were the result of drunk driving.[77] Of the roughly 10.4 million drunk driving arrests that are made,[78] approximately one third involve repeat offenders.[79] The cost of accidents alone to federal and local governments, and taxpayers, was estimated in 2002 (and thus, undoubtedly a conservative estimate of today's costs) to be approximately fifty-one billion dollars per year, excluding deaths and injuries.[80] Steven D. Levitt and J. Porter estimate the externality imposed on society by drunk driving may be as high as $8,000 for each incident of drunk driving.[81] Drunk driving, thus, has important negative consequences on society, and, like with drug use, involves sentencing for at least a segment of the population that is prone to addiction. Regulating crimes that stem from addictive behavior presents challenges in which increased punishment may be ineffective. Understanding empirically when subgroups are not responsive to more traditional sanctions may result in more effectively targeting sanctions, and reallocating resources effectively, to reduce the recurrence of the crime.

---

77. NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., TRAFFIC SAFETY FACTS: ALCOHOL-IMPAIRED DRIVING 6 tbl.4 (2010), www-nrd.nhtsa.dot.gov/Pubs/811606.pdf [hereinafter TRAFFIC SAFETY FACTS 2010].

78. In all likelihood, as Roth points out, this commonly cited figure from the FBI crime statistics is likely to be a conservative estimate. *See* RICHARD ROTH, 2013 SURVEY OF CURRENTLY-INSTALLED INTERLOCKS IN THE U.S. 2 n.2 (2013), http://www.transportation.nebraska.gov/nohs/pdf/IgnitionInterlockSurveyUS.pdf.

79. NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., CASE STUDIES OF IGNITION INTERLOCK PROGRAMS 5 (2012); Philip J. Cook & Maeve E. Gearing, *The Breathalyzer Behind the Wheel*, N.Y. TIMES (Aug. 31, 2009), http://www.nytimes.com/2009/08/31/opinion/31cook.html?_r=0; *see also* FBI, CRIME IN THE UNITED STATES 2011: ESTIMATED NUMBER OF ARRESTS tbl.29 (2011), http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/tables/table-29.

80. BLINCOE ET AL., *supra* note 15, at 2.

81. Steven D. Levitt & Jack Porter, *How Dangerous are Drinking Drivers?*, 109 J. POL. ECON. 1198, 1198 (2001).

IV.   THE EMPIRICAL APPROACH

A. The Regression Discontinuity Design

The "gold standard" for making causal inferences with quantitative methods is the randomized experiment.[82] In the case of sentence enhancements and DWI recidivism, the ideal experiment would involve randomly assigning the sentence either at the legal limit or an enhanced sentence (in this case, the treatment) to first-time offenders, while a control group would receive no sanctions or the sentence without the enhancement. This experience would include a large number of defendants in the treatment and control groups, so that in expectation, the process of random assignment would ensure that both the treatment and control groups would be comparable to each other on both observed and unobserved characteristics associated with the treatment.[83] While an experiment of this sort would be ideal for causal inference, a number of practical limitations, including ethical issues, make it unlikely to be completed.

Since sentence enhancements are not randomly assigned, I rely on a quasi-experimental design—the regression discontinuity design (RDD)—to make causal inferences about the effect of sanctions at the legal limit or an enhanced sentence on recidivism and vehicle accidents for individuals. The idea with the RDD is that there is a discontinuous threshold or cut-point that determines who receives a treatment. The technique was first used in a paper by the educational psychologists Donald Thistlethwaite and Donald Campbell,[84] who evaluated the effect of receiving a National Merit Scholarship on "attitudes toward intellectualism,"[85] success in obtaining college scholarships, and future academic and career performance.[86] Their research design provides a clear example of how the RDD works, which will be helpful in understanding how the technique can be used for this study. The authors compared groups of near winners of the National Merit Scholarship, with those who barely qualified for the scholarship.[87] The scholarship, which

---

82. Stephen G. West et al., *Alternatives to the Randomized Controlled Trial*, 98 AM. J. PUB. HEALTH 1359, 1359 (2008) (describing the well-established paradigm of randomized experiments).

83. We say "in expectation," because any one randomization can lead to imbalance on observable or unobservable characteristics between the treatment and control groups that are correlated with the treatment.

84. Donald L. Thistlethwaite & Donald T. Campbell, *Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment*, 51 J. EDUC. PSYCHOL. 309, 309–11 (1960).

85. *Id.* at 309.

86. *Id.*

87. *Id.* at 310.

in this case is the treatment, requires receiving a minimum score on the PSAT, a standardized test taken by most high-school students in the United States.[88] In a randomized experiment, with large numbers, random assignment establishes the expectation of equivalence between the treatment and control group. RDD, by contrast, relies on non-random assignment, where a known cut-off point in the assignment of the treatment creates a discontinuity in the receipt of treatment at that threshold.[89] In the case of the National Merit Scholarship, the minimum required score on the PSAT would be the forcing variable[90] that would determine the receipt of the treatment (in this case, a scholarship) that would create a discontinuity in the number of high school students that received the scholarship.[91] The key insight of Thistlethwaite and Campbell's paper, as the economist Wilbert van der Klaauw points out, is that one could use the group just below the cut-off (in the case the near-winners of the scholarship) as a comparison group for those who did receive the treatment.[92] The key assumption, which to an extent is statistically testable, is that the group below the cut-off is a valid comparison group with the group that receives the treatment. If those conditions are met, then the assignment near the threshold that triggers assignment of the treatment can be viewed as being "as-if random," thus enabling strong causal inferences to be made about the effect of the treatment on the outcome. Thus, with a high degree of confidence, Thistlethwaite and Campbell concluded, in comparing near-winners to winners of the National Merit Scholarship, that the scholarship increased the likelihood that the recipient would receive future scholarships, but the scholarship did not affect student attitudes toward education or career plans.[93]

---

88. *See* NAT'L MERIT SCHOLARSHIP CORP., OFFICIAL STUDENT GUIDE TO THE PSAT/NMSQT 4–5 (2015) http://www.nationalmerit.org/student_guide.pdf.

89. JOSHUA D. ANGRIST & JÖRN-STEFFEN PISCHKE, MOSTLY HARMLESS ECONOMETRICS 251 (2009). To be more precise, the cut-off point can either create a known cut-off point, or there can be a known threshold that increases the probability of receiving the treatment. The latter case is typically referred to as a fuzzy RDD, while the former is referred to as a sharp RDD.

90. Some scholars also refer to the forcing variable as the assignment variable. *See, e.g.*, David S. Lee & Thomas Lemieux, *Regression Discontinuity Designs in Economics*, 48 J. ECON. LITERATURE 281, 281 (2009).

91. For a more extensive discussion of RDD, see generally ANGRIST & PISCHKE, *supra* note 89; Devin Caughey & Jasjeet S. Sekhon, *Elections and the Regression Discontinuity Design: Lessons from Close U.S. House Races, 1942–2008*, 19 POL. ANALYSIS 385 (2011); Guido W. Imbens & Thomas Lemieux, *Regression Discontinuity Designs: A Guide to Practice*, 142 J. ECONOMETRICS 615 (2008); Lee & Lemieux, *supra* note 90; Guido W. Imbens & Jeffrey M. Wooldridge, *Recent Developments in the Econometrics of Program Evaluation*, 47 J. ECON. LITERATURE 5 (2009); Wilbert van der Klaauw, *Regression-Discontinuity Analysis: A Survey of Recent Developments in Economics*, 22 LABOUR 219 (2008).

92. van der Klaauw, *supra* note 91, at 220.

93. Thistlethwaite & Campbell, *supra* note 84, at 317.

In the case of sentence enhancements, I take advantage of exogenous thresholds in DWI laws to make causal inferences about their effects on recidivism and subsequent vehicle crashes. The RDD compares defendants at various blood alcohol content (BAC) levels (the forcing variable), which determine if sanctions are administered at all, or if an enhancement is given. Of particular interest is the legal limit, which starting in most states in 2002, was a BAC level of 0.08.[94] In addition, a BAC of 0.15 triggers an increased license suspension from 120 days to 180 days.[95] I look at the effect of the sentence on the defendant's propensity to recidivate and have vehicle crashes following their first offense. With large comparison groups just below and just at or above the BAC cut-off for the legal limit or the enhanced sentence, one can compare these groups in a similar manner to the comparison groups in the Thistlethwaite and Campbell RDD. Since it is theoretically likely (and statistically testable with regard to observable pre-treatment variables) that the BAC levels in these two different groups are non-strategically chosen, the discontinuity specification allows for the treatment assignment to be "as-if random," as was the case with Thistlethwaite and Campbell's RDD.[96] Appendix I(D) describes the model and estimation strategy in greater detail.

Although the method received little notice when first introduced by Thistlethwaite and Campbell in 1960, use of the technique has experienced immense growth in the last decade starting with a series of papers that examined the effect of financial aid given on student enrollment decisions,[97] and the effect of class size on student achievement.[98] Scholars have also used the method in a number of different domains in the study of crime.[99] To date,

---

94. TRAFFIC SAFETY FACTS 2010, *supra* note 77, at 1.

95. *See infra* tbl.1.

96. Thistlethwaite & Campbell, *supra* note 84, at 317.

97. Wilbert van der Klaauw, *Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach*, 43 INT'L ECON. REV. 1249, 1249 (2002); Wilbert van der Klaauw, *A Regression-Discontinuity Evaluation of the Effect of Financial Aid Offers on College Enrollment* (C.V. Starr Center for Applied Economics, Working Paper 97-10, 1997), http://ideas.repec.org/p/cvs/starer/97-10.html.

98. Joshua D. Angrist & Victor Lavy, *Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement*, 114 Q.J. ECON. 533, 535–36 (1999).

99. *See, e.g.*, Richard A. Berk & Jan de Leeuw, *An Evaluation of California's Inmate Classification System Using a Generalized Regression Discontinuity Design*, 94 J. AM. STAT. ASSOC. 1045, 1045 (1999); Richard A. Berk & David Rauma, *Capitalizing on Nonrandom Assignment to Treatments: A Regression-Discontinuity Evaluation of a Crime-Control Program*, 78 J. AM. STAT. ASSOC. 21, 21 (1983); M. Keith Chen & Jesse M. Shapiro, *Do Harsher Prison Conditions Reduce Recidivism? A Discontinuity-Based Approach*, 9 AM. L. ECON. REV. 1, 2 (2007); Lee & McCrary, *supra* note 25, at 1–4; Randi Pintoff, *The Impact of Incarceration on Juvenile Crime: A Regression Discontinuity Approach* 2 (Job Market Paper, 2004), http://economics.yale.edu/sites/default/files/files/Workshops-Seminars/Labor-Public/pintoff-041008.pdf.

with two exceptions, I am unaware of anyone who has used BAC levels to examine the causal effect of sentencing enhancements on recidivism. Ian Ayres may have been the first to publish about the possibility, in an example mentioned in a tribute to law and economics scholar Thomas Ulen.[100] Economist Benjamin Hansen recently examined the effect of drunk driving enhancements on recidivism in Washington State using the regression discontinuity approach.[101] One particularly noteworthy experimental study is Martin, Annan, and Forst's 1993 study, which exploits random assignment of 383 defendants convicted of drunk driving to one "harsh" and one "lenient" judge in Minnesota to determine whether harsher sentences have an effect on recidivism.[102] The authors found no statistically significant difference in the recidivism rates of persons sentenced by judges who tended to incarcerate defendants more often from those who did so less frequently.[103]

This study takes advantage of a setting where the institutional conditions enable strong causal inferences to be made from the regression discontinuity design. Because strategic sorting around the discontinuity can undermine the causal inferences made about the effect of the sentence enhancement, the paper focuses on Arkansas, a location where police, prosecutorial, and judicial discretion is extremely limited. Specifically, in this state: (1) there is electronic reporting of BAC results, making it difficult to under- or over-report results; (2) there is no charge bargaining or plea bargaining for DWI; and (3) judges are not able to expunge DWI offenses for the time period of interest.[104] Arkansas offers all of these conditions, and because the criminal code prohibits the expungement of DWI and other crimes, which can threaten a research design's external validity and interpretation of the results of the treatment, I believe the state is a nearly ideal location to conduct this study. With previously untapped micro-data, the research provides insight not only into the effectiveness of various punishments, but also permits the descriptive examination of court processes and the dynamics of prosecutorial and judicial decision-making.

---

100. Ian Ayres, *Very Like a Law Professor: An Essay in Honor of Tom Ulen*, 2011 U. ILL. L. REV. 1767, 1774–75 (2011).

101. Benjamin Hansen, *Punishment and Deterrence: Evidence from Drunk Driving*, 105 AM. ECON. REV. 1581, 1581–82 (2015).

102. Susan E. Martin et. al., *The Special Deterrent Effects of a Jail Sanction on First-Time Drunk Drivers: A Quasi-Experimental Study*, 25 ACCIDENT ANALYSIS PREVENTION 561, 562–63 (1993).

103. *Id.* at 563.

104. These institutional features are discussed in greater detail *infra* Section V.

B. *Difference-in-Differences Estimation*

While the RDD helps rigorously identify the causal effect of punishment within a time period, understanding the effect of laws across time periods requires an approach that can help account for differences related to time that might be driving the results. In order to examine the effect of legal changes to the drunk driving regime in Arkansas, I rely on a difference-in-differences (DID) approach, which allows for the estimation of the causal effect of the law on recidivism and subsequent vehicle accidents.

DID relies on an experimental framework that allows one to make causal inferences about the effect of changes in drunk driving laws over time.[105] If we simply compare the RDD estimate in one legal regime to the estimate in another period, we might worry that the individuals under one legal regime might be affected by temporal trends, or that discrete events in time unrelated to the treatment (in this case, drunk driving sanctions at the legal limit or enhanced sanctions implemented by a new law or an amendment to an existing law) might affect the results.[106] As a result, the construction of a comparison group that stretches across both time periods that was not exposed to the treatment can be used to account for temporal trends in the outcome that are not the result of being exposed to the treatment.

In the case of drunk driving sanctions, I examine three important changes to the legal regime for drunk driving in Arkansas. These include (1) the revocation of a restricted permit allowing DWI offenders with BACs greater than or equal to 0.15 to drive to work, school, and for a few other purposes in 2003; (2) the shift in control of IIDs from the courts to the Office of Driver Control in 2005; and (3) an increase in license suspension length, the revocation of restricted permits for all DWI offenders, and mandatory attendance at a victim's panel in 2009.[107] I examine the effects of these changes on recidivism and subsequent vehicle crashes before and after each legal reform—the "first difference." In order to have a treatment and control group spanning the entire period, I compare the group that is just at or above the legal BAC limit with those that are just below. I make the same comparison for those at or just above the enhanced BAC level with those that

---

105. For more in-depth detail on DID estimation, see WILLIAM H. GREENE, ECONOMETRIC ANALYSIS 155–58 (Rod Banister et al. eds., 7th ed. 2011); JEFFREY M. WOOLDRIDGE, ECONOMETRIC ANALYSIS OF CROSS SECTION AND PANEL DATA (2nd ed. 2002); Alberto Abadie, *Semiparametric Difference-in-Differences Estimators*, 72 REV. ECON. STUD. 1 (2005); Marianne Bertrand et al., *How Much Should We Trust Differences-in-Differences Estimates?*, 119 Q. J. ECON. 249 (2004); Guido Imbens & Jeffrey M. Wooldridge, Presentation at the National Bureau of Economic Research Summer Institute 2007: Difference-in-Differences Estimation (July 31, 2007), http://www.nber.org/minicourse3.html.

106. Abadie, *supra* note 105, at 1.

107. These reforms to the drunk driving regime are discussed in detail *supra* Section V.

are just below the 0.15 BAC threshold. A more formal description of the model, with an in-depth discussion of the assumptions, is available in Appendix I(D).

One important assumption made with the DID approach merits discussion with respect to this particular study. The control group is assumed to have followed a similar trend to the treatment group, except for the addition of the treatment. What is important to note is that a similar time trend does not mean that the mean outcome has to be the same for a given time period; rather, the two trends follow each other, even if at different levels. To a certain extent, this assumption is made more acceptable by examining whether there are pre-treatment differences in characteristics between the treatment and control groups. However, some unobserved policy change or other variable that affects both groups at the time of the law's passage would undermine the integrity of the counterfactual. Since Arkansas has a part-time legislature that meets infrequently, analyzing all of the laws passed by each session could be completed relatively easily. I investigated legislation that passed in the similar session, and found no piece of contemporaneous legislation that would likely affect the results.[108]

C. Interviews and Police Reports

In addition to the quantitative work, I conducted numerous in-depth interviews with relevant actors in the criminal justice and political systems. These interviews served a number of purposes. First, the interviews provided important institutional context in order to understand aspects of the criminal justice system in Arkansas. For example, while the legislative design limited police, prosecutorial, and judicial discretion, knowing ways of potentially circumventing that system was important. Second, interviews provided key information about mechanisms that shed light on how sentencing policy influenced reductions (or increases in recidivism).

In addition to interviews, I also obtained police reports and data from local jurisdictions in the state. The reports gave helpful context in terms of how the process worked for booking an adult with DWI, and also provided information on one important aspect of police discretion: the time law enforcement officers could give an individual between their initial arrival on the scene and when the person takes a court-admissible alcohol test. In addition to giving qualitative texture to the process, police records have also

---

108. I also include year fixed effects (dummy variables for each year), which account for the passage of legislation that might be correlated with the treatment in a specific a year.

been used to do out-of-sample verification of the integrity of data I have obtained from other sources.

V.   CASE SELECTION: THE ARKANSAS STATUTORY AND ENFORCEMENT REGIME FOR DRUNK DRIVING

Arkansas is a nearly ideal setting to conduct this analysis because of a combination of its statutory regime, data quality and availability, and administrative procedures. In order to isolate the causal effect of the enhancement on recidivism, the existence of limited police, prosecutorial, and judicial discretion enables reliable causal inferences without imposing strong assumptions on the data.

A. Police Discretion

Police discretion in the form of under- or over-reporting of BAC levels would undermine the reliability of causal inferences made about the effect of the sanctions, since officers would likely be inaccurately reporting on the basis of characteristics unobservable to the researcher. In Arkansas, breath test results are immediately and automatically reported electronically to the State Health Department's Office of Alcohol Testing (OAT), and in most cases, to the local police department. OAT is responsible for calibrating breathalyzer equipment, establishing standards for and certifying acceptable equipment, training personnel in using the equipment, and maintaining BAC records. The records are independently maintained by the agency, and courts frequently rely upon the records when making sentencing decisions. OAT also examines blood and urine tests when someone is suspected of alcohol intoxication (DWI). Blood and urine tests are typically administered when a driver: (1) is too intoxicated to perform a breath test; (2) is incapacitated as a result of an accident; or (3) is involved in a serious accident where major bodily harm or death has occurred. Blood and urine tests, whether done at local hospitals or at a police station, are also reported to OAT when alcohol is suspected. OAT has data on approximately 25,000 individuals per year offered breath tests throughout the state, and some 1,500 blood and urine tests per year. Because the paper focuses on the effect of sentencing on recidivism, the treatment (in this case, the sanctions close to the legal limit or the sentence enhancement) is conditioned on a potential offender being arrested for DWI. Therefore, police discretion in terms of who gets arrested should not affect the analysis of the causal impact of sanctions on recidivism.

While discretion in BAC reporting is unlikely to be taking place, the police can exercise discretion in the timing of administration of the breathalyzer

tests. Typically, upon suspicion of DWI, an officer performs field sobriety tests on the driver upon stopping the vehicle. An officer can also administer a portable breathalyzer test, but the results are not admissible in court. During the time period I study in the jurisdictions from which I obtained data, officers are required to perform two alcohol tests, which could be any combination of a field sobriety, breathalyzer, urine or blood test. While blood tests are typically done at local clinics, the other tests are typically performed at the site of the stop and the police station. By statute, during the time period of study, the lower of the two BAC results should be counted for adjudication. In addition, I conducted interviews to determine whether this strategic behavior takes place with respect to the timing of breath tests, and if so, how officers selectively manipulate the timing of BAC testing. Because all test results are reported electronically, and they are used in court proceedings, I believe that safeguards in the form of monitoring are in place to prevent this form of manipulation from happening.

B. *Prosecutorial Discretion*

In addition to the under- or over-reporting of alcohol test results, prosecutorial discretion could also undermine the causal inference strategy of the research design. For instance, if prosecutors selectively and systematically charge defendants with an offense below their BAC level in a manner that is correlated with unobserved factors, the causal inferences made from the quasi-experimental nature of the design are likely to be undermined. In addition, dropped cases based on unobservable factors could undermine the random assignment of cases in the neighborhood of the discontinuity.

In Arkansas, *every* DWI case where an arrest charge took place must be prosecuted. Specifically, section 5-65-107(a) of the Arkansas DWI Omnibus Act states that any person arrested for DWI "shall be tried on those charges or plead to those charges, and no such charges shall be reduced or dismissed."[109] The constitutionality of this provision based on violation of separation of powers and the doctrine of prosecutorial discretion emerged as the central issue in a number of court cases. In *Sparrow v. State*, the court ruled it was not unconstitutional for the Omnibus DWI Act to (1) mandate prosecution of the arrest charge, (2) prohibit charge bargaining, and (3) disallow plea bargains.[110] Similarly, in *Southern v. State*, the court held the

---

109. ARK. CODE ANN. § 5-65-107(a) (2015) (originally passed in 1983).

110. 683 S.W.2d 218, 219 (Ark. 1985). Specifically, the court rejected the argument "that the DWI law violates the separation of powers provision in the Arkansas and United States Constitutions in that it takes away from the prosecuting attorney and the court the right to reduce

"doctrines of prosecutorial discretion and separation of powers are [not] violated by" this section.[111] The court reiterated this holding in *Johnston v. City of Fort Smith*.[112] Finally, in *Bigham v. State*, the court held that "[i]t is not unconstitutional for this act to authorize a police officer, rather than the prosecuting attorney or grand jury, to file the misdemeanor charge."[113] Interviews with prosecutors, judges, court clerks, and police, along with examination of the data in a number of jurisdictions reveal that this law is closely followed in practice. One court clerk stated that Mothers Against Drunk Driving (MADD) kept a close watch on the court where she works to make sure that every case was adjudicated and no plea bargains were being made for DWI defendants.[114] Based on interviews and analysis of the data in seven jurisdictions, charges tend only to be amended when the court is unable to find definitive evidence that a prior DWI took place.[115] Moreover, prosecutors are permitted to drop DWI cases only in two rare instances: when an accused offender has not been arraigned within one year of the offense, or when identity theft has taken place.[116] By statute, the prosecutor is supposed to drop cases after a warrant is issued and the person has not been found within one year of the arrest date. In practice, in the jurisdictions I examined, the court keeps these records for more than a year, and they are periodically dropped after a few years. Though records are still kept in the court database, no cases in the jurisdictions I examined were prosecuted if the defendant was not found for more than one year.

C. Judicial Discretion

Judicial discretion varies with respect to different penalties for DWI in Arkansas. Presently, six forms of sanctions are possible for those arrested for DWI: (1) license suspension, (2) incarceration, (3) fines, (4) safety school, (5) rehabilitation, and (6) ignition interlock devices. Figure 2 shows the DWI penalty regime that has been in place since 2001, much of which remains in place from the original passage of the 1983 DWI Omnibus Act, the main piece of legislation that governs procedural and sentencing regime for drunk driving in Arkansas.

---

a charge and accept plea bargains and places that power within the hands of the policeman, who files the charge." *Id.*

    111. 683 S.W.2d 933, 934–35 (Ark. 1985).

    112. 690 S.W.2d 358, 361 (Ark. Ct. App. 1985).

    113. 743 S.W.2d 405, 406 (Ark. Ct. App. 1988).

    114. Telephone Interview with unnamed court clerk, Garland Cty. Dist. Ct. (Apr. 4, 2011).

    115. *Id.*

    116. *Id.*; Telephone Interview with unnamed court clerk, Garland Cty. Dist. Ct. (Oct. 19, 2011).

Table 1 shows the DWI penalty regime that was in place from 2001 until 2013. Although sentencing guidelines in Arkansas are voluntary and judges are allowed to depart from the guidelines in "non-typical" cases without written justification for the departure,[117] in the cases I examined, I found no evidence of a departure from the statewide voluntary sentencing guideline regime.

117. Neal B. Kauder & Brian J. Ostrom, *State Sentencing Guidelines: Profiles and Continuum*, NAT'L. CENT. FOR STATE CTS. (July 2008), http://www.ncsc.org/~/media/Microsites/Files/CSI/State_Sentencing_Guidelines.ashx.

Table 1: The Regulatory Regime for Drunk Driving in Arkansas 2001–2013

Sanction	2001–2003	2003–2005	2005–2009	2009–2013
First-Time Offenders				
License Suspension	BAC [0.08, 0.15): 120 day license suspension with restricted permit	BAC [0.08, 0.15): 120 day license suspension with restricted permit	Same as 2003–2005	BAC $\geq$ 0.08: 180 day license suspension with no restricted permit
	BAC $\geq$ 0.15: 180 day license suspension with restricted permit	BAC $\geq$ 0.15: 180 day license suspension with no restricted permit	Same as 2003–2005	
Interlock	Interlock under jurisdiction of courts	Same as 2001–2003	Interlock given to Driver Control	Same as 2005–2009
Rehabilitation				BAC $\geq$ 0.08: Mandatory Attendance at a Victims Panel
Incarceration	1 day–1 year	1 day–1 year	1 day–1 year	1 day–1 year
Fines	$500–$5000	$500–$5000	$500–$5000	$500–$5000

Judges have greater discretion with incarceration, fines, safety school, rehabilitation, and IIDs than with license suspensions. This increased relative discretion exists in large part because license suspensions were largely taken

out of the hands of the judiciary, and handled by the Office of Driver Control starting in 2005.

D. *License Suspensions*

License suspensions, together with incarceration and fines, have long been the main penalties in place for drunk driving in the United States.[118] In Arkansas, Act 1501 and Act 5601, both passed in 2001, lowered the BAC level for a 180-day license suspension from 0.18 and above to 0.15 and above.[119] During this time, those with a suspended license could still obtain a restricted permit to drive to and from home to work. However, in 2003, the state legislature passed Act 1779, which took away the restricted permit for first-time offenders with a BAC of 0.15 and above.[120] In 2009, the legislature passed Act 1293, which eliminated the license suspension enhancement; all first-time offenders, regardless of their BAC, were given a six-month license suspension and no restricted permits could be issued.[121]

The formal enhancement only exists for license suspensions for first-time offenders. Judges have relatively limited control over license suspensions. Once a driver is arrested for drunk driving, the Office of Driver Control issues a temporary license to the defendant that is valid for 30 days. The defendant then has seven days to contest the suspension. If the defendant does not win the appeal, a license suspension goes into effect on the thirty-first day. Judges rarely, if ever, intervene in the license suspension process. With only rare exceptions, license suspensions are terminated only if the defendant is determined to be not guilty by the court. Thus, judicial discretion is relatively limited with this form of punishment. As additional court data becomes available, I will test to see if this pattern holds up, and will investigate the reasons for this variation both quantitatively and qualitatively.

---

118. *See, e.g.*, Elder et al., *supra* note 75, at 362 ("For the first two thirds of the 20th century, the traditional penalties assessed for a DWI conviction were jail, fines, and license suspension.").

119. To Amend Suspension and Revocation of Driving Privileges, sec. 1, § (a)(i), 2001 Ark. Legis. Serv. 1501 (West).

120. Act 1779 of 2003, No. 2716, 2003 Ark. Acts, http://www.arkleg.state.ar.us/assembly/2003/R/Acts/Act1779.pdf.

121. Act 1293 of the Regular Session, No. 1640, 2009 Ark. Acts, http://www.arkleg.state.ar.us/assembly/2009/R/Acts/Act1293.pdf.

E. Ignition Interlock Devices

As of 2011, all fifty states in the United States have IID laws in place as part of their sanctions regime for drunk driving.[122] Although a number of studies have shown that the devices are incredibly effective, with reductions in DWI recidivism ranging from 50 to 90 percent while IIDs are installed in the vehicle,[123] recent survey estimates from June 2013 show only that about 300,000 to 325,000 IIDs are in use, compared to an estimated 10.4 million arrested for driving under the influence.[124]

In 2005, the Arkansas legislature removed jurisdiction over IIDs from the courts, and gave it to an administrative agency. This move led to an increase in the number of interlock devices, and also provides an opportunity to examine how courts versus agency administer punishment. Mark Kleiman, in his seminal book, *When Brute Force Fails*,[125] discusses the importance of delivering punishment in a swift and certain manner. One possibility that led to the uptake is the lack of discretion that the administrative agency had, relative to judges, in disseminating IIDs. IIDs have always been voluntary in Arkansas, but they were made the only legal way to drive for first-time offenders in 2009. The number of IIDs surged in that year, and details of this reform will be discussed in greater detail later in this paper.

F. Expungement

Since the passage of the Community Punishment Act ("Acts 548 and 549" of the Arkansas Criminal Code) in 1993,[126] jurisdiction over expungement in

---

122. NHTSA REPORT, *supra* note 64, at 25–32 app.B. In 2009, 47 of the 50 states, including the District of Columbia, had IID laws in place (Alabama, South Dakota, and Vermont were the three exceptions). In 2011, Alabama became the fiftieth state to enact IID laws as part of its sanctions regime for drunk driving. *Ala. Becomes 50th State to Enact Ignition Interlock Legislation*, 12 WFSA (June 3, 2011, 4:54 PM), http://www.wsfa.com/Global/story.asp?S=14837540.

123. *See, e.g.*, NHTSA REPORT, *supra* note 64, at 22; Jeffrey H. Coben & Gregory L. Larkin, *Effectiveness of Ignition Interlock Devices in Reducing Drunk Driving Recidivism*, 16 AM. J. PREVENTIVE MED. 81, 81 (1999); A. Scott Tippetts & Robert B. Voas, *The Effectiveness of the West Virginia Interlock Program on Second Drunk-Driving Offenders*, 14 INT'L COUNCIL ON ALCOHOL, DRUGS & TRAFFIC SAFETY REP. 185, 189 tbl.2 (1997); L. Vézina, *The Québec Alcohol Ignition Interlock Program: Impact on Recidivism and Crashes*, 16 INT'L COUNCIL ON ALCOHOL, DRUGS & TRAFFIC SAFETY REP. 97 (2002); Robert B. Voas & Paul R. Marques, *Barriers to Interlock Implementation*, 4 TRAFFIC INJURY PREVENTION 183, 183 (2003); Charlene Willis et. al., *Alcohol Ignition Interlock Programmes for Reducing Drink Driving Recidivism*, 3 COCHRANE LIBR. 4168 (2009).

124. ROTH, *supra* note 78, at 2 fig.1.

125. MARK A. R. KLEIMAN, WHEN BRUTE FORCE FAILS: HOW TO HAVE LESS CRIME AND LESS PUNISHMENT 3 (2010).

126. ARK. CODE ANN. § 16-93-1203 (2005).

Arkansas has remained largely with the courts. Prior to the passage of the Community Punishment Act,[127] the Parole Board had the ability to expunge offenses. Under section 16-90-902 of the Arkansas Criminal Code, an individual who has been granted an expungement is permitted to state that the offense never occurred and that no record exists for the offense.[128] However, although as a matter of law the offense is viewed as never having taken place, under section 16-90-901 of the same code—with only a few exceptions—an expungement does not result in the physical destruction of any records.[129] In practice, expungement without the destruction or deletion of records is more akin to a case being sealed, rather than expunged in the traditional sense. One exception where the destruction of records is permitted is when no guilty verdict occurs. In this instance, Arkansas Criminal Code section 16-90-901 allows for the expungement of documents such as arrest records, orders, docket sheets, and any other case-specific documents.[130]

Arkansas law prohibited expungement for DWI until July of 2011.[131] In July 2011, a law went into effect making expungement possible five years after the DWI offense took place. Specifically, section 5-65-108(c)(1) of the DWI Omnibus Act prohibited the expungement of records for any defendant charged with a DWI offense.[132]

---

127. *Id.*

128. ARK. CODE ANN. § 16-90-902 (2013), *repealed by* Act 1460 of the Regular Session, No. 1638, 2013 Ark. Acts, http://www.arkleg.state.ar.us/assembly/2013/2013R/Acts/Act1460.pdf.

129. ARK. CODE ANN. § 16-90-901 (2013), *repealed by* Act 1460 of the Regular Session, No. 1638, 2013 Ark. Acts, http://www.arkleg.state.ar.us/assembly/2013/2013R/Acts/Act1460.pdf.

130. *Id.*

131. Act 626 of the Regular Session, No. 1608, 2011 Ark. Acts, http://www.arkleg.state.ar.us/assembly/2011/2011R/Acts/Act626.pdf.

132. Other crimes vary in terms of whether or not they can be expunged. Act 1035, which went into effect in 1999, made certain drug offenses eligible for expungement, while also prohibiting expungement for Murder II, Manslaughter, Negligent Homicide, Sexual Abuse I, Battery I, Domestic Battery I, and Simultaneous Possession of Drugs and Firearms. However, the expungement prohibition can be time limited for some offenses. For example, domestic battery is eligible for expungement five years after the offense took place. There are six cases where individuals are typically eligible for expungement. These cases merit attention because the impact of expungement on the most important pre-treatment covariate—criminal history—is important to keep in mind when analyzing covariate balance, which is discussed in Section 50.2. These six cases include: (1) offenders who are pardoned, except those pardoned for offenses that: (a) involved offenses against minors; (b) resulted in death or serious injury; and (c) sex offenses. ARK. CODE ANN. § 16-90-605 (1995), *repealed by* Act 1460, H.B. 1638, 89th Gen. Assemb., Gen. Sess. (Ark. 2013); (2) first-time offenses involving most criminal, driving (although not DWI), and controlled substance cases where the convicted individual successfully completes all probation terms. ARK. CODE ANN. § 5-65-108 (2015); (3) minors who were pardoned for offenses they committed when they were under 16 years old. ARK. CODE ANN. § 16-90-601 (1995), *repealed by* Act 1460 of the Regular Session, No. 1638, 2013 Ark. Acts; (4) individuals convicted of a non-violent felony if the act took place while the person was under 18 years old. ARK. CODE ANN. § 16-90-602 (1995), *repealed by* Act 1460 of the Regular Session, No. 1638, 2013 Ark.

The expungement prohibitions feature of the Arkansas sentencing regime not only increase the precision with which estimates of recidivism are made, but also allows for the use of the entire sample of cases, thus increasing the external validity of claims made about the population of interest.

G. *Strategic Behavior and Other Threats*

Because RDDs can be invalidated if individuals are able to precisely manipulate the forcing variable, a discussion of potential forms of strategic behavior by defendants is in order. Strategic sorting and manipulation around the threshold is especially important with sentencing, since defendants at or just above the BAC level that triggers the sanctions have strong incentives to lower their BAC level so they can increase their chances of receiving a reduced punishment.

Although drivers and police are likely to have a sense of the general range of their individual BAC, cases close to the threshold triggering sanctions at the legal limit or an enhanced penalty may have a number of factors giving drivers some control over their BAC levels. First, drivers could carry their own breathalyzers, since they are available to individual consumers. While additional empirical work will be conducted to verify the extent of its presence in the population of interest, I believe that this is unlikely to constitute a significant percentage of drivers in the sample. Even if drivers are carrying their own personal breathalyzers, these are very prone to measurement error, especially within the BAC threshold ranges that I am studying. Second, while an individual's weight is a relatively good predictor of one's BAC, variation still exists in individual responses to alcohol. These sources of uncertainty undermine the control that any given individual may have over his or her precise BAC level (in this case to the hundredth of a decimal place). Third, though a challenging process, individuals could strategically manipulate their identity around the discontinuity. This possibility especially exists for undocumented immigrants, where law enforcement agencies find difficulty in maintaining consistent records. Similarly, identity theft may also result in measurement error. Interviews with

---

Acts; (5) a person who (a) successfully completes probation or has an expungement-eligible offense or (b) successfully completes a commitment to the Department of Corrections or Department of Community Correction and who (i) has one or no felony convictions that were not for a capital offense, first or second degree murder, first degree rape, kidnapping, aggravated robbery, or delivery of controlled substances to a minor or who (ii) has no prior felonies. ARK. CODE ANN. § 16-93-1207 (2015); and (6) any individual who is charged and arrested for any criminal offense and is *nolle prossed*, dismissed, or acquitted. ARK. CODE ANN. § 16-93-1207 (2015).

police, judges, and city attorneys in the area confirm that this is an issue; precise estimates are currently unknown, but the likelihood of it affecting estimates disproportionately across the BAC threshold are very low. Finally, most drivers are unlikely to know exactly how many drinks—to a precise level—will move them to specific BAC levels. Appendix I(B) shows the number of drinks an individual would likely have to be at a certain BAC level, given an individual's sex and weight. For most people within normal weight ranges, one drink can result in movement across a wide range of the BAC scale, also adding to the difficulty of precisely landing just below, or just at or above a threshold that triggers a particular sentence.

VI.   DESCRIPTION OF DATA

Against this backdrop, I obtained data from a number of government agencies in Arkansas. Specifically, I have alcohol testing and vehicle accident data from OAT, court data from the Administrative Office of the Courts and a number of local courts, and local police data.

A. Alcohol Testing Data

BAC data came from OAT in the form of individual-level reports. OAT provided me with every blood, breath, and urine test that was given by an official health worker or law enforcement officer in the state. All BACs— whether they are blood, breath, or urine—are to the thousandth of a decimal place. Sentencing for drunk driving is done at the hundredth of a decimal place, where all digits in the thousandths place are rounded down (e.g. a BAC of 0.089 is classified as 0.08).[133] The time series for analysis runs from March 6, 2001 through June 11, 2013.[134] During this period, officials administered 186,745 alcohol tests for sentencing in the state.[135] These alcohol tests include tests given in health clinics, hospitals, jails, morgues, and police stations. As a consequence of the Health Insurance Portability and Accountability Act (HIPAA), tests given by a private practitioner are not included. In an email exchange, the OAT Director stated that these tests "over the past few years"

---

133. As a result, specifications are presented with robust standard errors. Results are also available from the author with standard errors clustered at the hundredths of a decimal place.

134. The data set includes alcohol tests from before March 6, 2001, but the OAT Director cautioned us against reliability of the data prior to this date, since information technology systems and practices were not as standardized, potentially undermining the reliability of the data. These observations are not used in the data analysis.

135. The caveat that these tests are used for sentencing is included, because the figure does not include repeat tests, given as a result of machine malfunction or test administration difficulties.

constitute about 3.25 percent of the overall sample.[136] The tests also include non-drivers, since passengers and pedestrians are sometimes tested, which are excluded from the analysis.

During the study period, the state used two types of breath machines. Until January 2009, every county used DataMaster branded machines. Starting on January 8, 2009, OAT implemented a phased rollout of Intoximeter breathalyzer machines; by March 25, 2011, DataMaster machines were no longer in use. DataMaster, blood, and urine test data contain the machine serial number; incident date; incident time; the time the observation period began; the location of the test; the arresting authority; the name and ID number of the operator; the subject's name; whether the subject was a driver, passenger, or pedestrian; whether an accident, injury, or fatality took place; testing indicators; the time the sample was taken; the test date and time; and the BAC testing results. If more than one sample was taken, the statute states that the sample with the lower result must be counted for sentencing. Intoximeter reports contain the same variables, and more detailed alcohol testing data. Appendix I (C) shows a redacted Intoximeter breath report, along with a report from a blood test. DataMaster output comes in the same exact form as the blood test.

In studying recidivism, constructing an accurate individual-level identifier is important, so that the effects of sentencing can be accurately determined. I submitted a number of Arkansas Freedom of Information Act (FOIA) requests, and eventually obtained, in addition to the individuals' names, their birth month, birth year, state of their driver's license, and the last four digits of their driver's license number, in order to create a person-level identifier. When the State of Arkansas used DataMaster machines, law enforcement officers and medical personnel entered the individual's information manually. As a result, data entry errors are possible, which would likely result in recidivism rates and treatment effects being understated, since the probability of categorizing the same person as two different people is more likely with the previously described procedures. I have no reason to believe that data entry errors would vary across the BAC thresholds that I explore, in large part because the person using the breath machine enters the individual's information before knowing the arrestee's breath test result.[137] Nevertheless,

---

136. By law, every law enforcement agency must submit samples to OAT, but in practice, alcohol tests might occasionally be conducted by a hospital. The number of tests that fall into this category are a small percentage of tests. Email from Laura Bailey, Chief Admin. Officer of the Blood Alcohol Testing Program, Ark. Dep't of Health, to author (Sept. 27, 2013, 2:32 PM) (on file with author).

137. If data entry was done after the BAC result was given, one possible explanation for higher error rates below the legal limit might be that the person entering the information did not

I took a number of steps to improve the accuracy of the identifier. First, I looked for obvious reporting errors through our own inspection of the data. Second, I used out-of-sample data from state and local courts, the Arkansas State Police crash database, local prosecutor data, and the Office of Drive Control to detect and correct any possible data entry errors. Third, I constructed multiple individual-level identifiers, using permutations of the identified information, and performed robustness checks on the results of our data analysis. I find that the results are substantively robust, irrespective of the identifier that is chosen.

B. *Vehicle Crash and Local Court Data*

Vehicle crash data came from the OAT alcohol testing reports and the Arkansas State Police crash database. The alcohol testing reports contain data on whether an individual who was tested was involved in a vehicle crash, and whether injuries or deaths resulted from the intoxication or crash.

In addition to statewide vehicle crash data, I obtained local court data from local counties and cities. This data includes demographic data, criminal history, sentencing reports, and court narratives of everyone at the court. In some cases, I had direct access to the court database that helped us understand and collect important data for this project.

C. *Descriptive Statistics*

I begin by examining the frequency of BAC test results, in order to see whether the distribution is continuous and smooth across the thresholds of interest. OAT measures BAC to the thousandths of a decimal place (3 decimal places), and so the data reflects the true measurement of the BAC tests. Any non-random sorting across the discontinuity can undermine the integrity of the RDD, since randomness local to the threshold is an important condition for making valid causal inferences about the effect of sanctions on the outcomes of interest.

Figure 2 presents histograms that show the frequency distributions of BAC tests during the entire time series of reliable data (March 2001 to July 2013). Figure 2(a) shows BAC counts for a wide BAC range that includes all BAC tests greater than 0 and less than or equal to 0.45. Figure 2(b) shows the frequency distributions for a narrower range of BAC levels between 0.05 and 0.20, which are closer to the thresholds of the legal limit of 0.08 and the

---

think the test results would be as consequential, and would thus be less diligent in performing data entry. We have no reason to believe this is the case as a consequence of the process.

license suspension enhancement of 0.15, which are the focus of this paper. If police officers or other actors systematically under- or over-reported BACs at the threshold, or if individuals could systematically manipulate their BAC levels strategically at the threshold on a large scale, a discontinuous "jump" at 0.08 or 0.15 would be visible. The histograms provide *prima facie* evidence that this type of behavior is unlikely because the frequency counts do not have sudden changes across both thresholds.

Figure 2: Frequency Distribution of BAC (Bin Width 0.001) for First-Time and All Offenders

Figure 2(a): Frequency Distribution for $0 < \text{BAC} \leq 0.45$

Figure 2(b): Frequency Distribution for $0.05 \leq \text{BAC} \leq 0.20$

D. *Similarity of Groups Across BAC Thresholds*

In order to make causal inferences about the effect of the sanctions being tested, one important assumption is that the group below the threshold is a

valid counterfactual for the group above the threshold.[138] To test the extent to which this assumption might be true, I examine whether the comparison groups just below and just at or above the threshold are similar with respect to pre-treatment characteristics. Specifically, I examine whether the groups at or above the BAC threshold of 0.15 are statistically indistinguishable on pre-existing (or pre-treatment) characteristics from those who are not (in statistical parlance, I see if the groups are balanced on pre-treatment covariates). I also check for covariate balance at the legal limit as well.

Figure 3 presents the covariate balance results for 37 pre-treatment variables thresholds close to the 0.08 ($0.06 \leq$ BAC < 0.10) and 0.15 ($0.13 \leq$ BAC < 0.17) thresholds. The data in Figure 3 comes from Garland County, where I obtained individual-level demographic and criminal history data from the District Court. Demographic covariates include age, height, weight, sex, and race. I show the results for the 32 most common offenses across seven different categories of offenses: (1) traffic violations, (2) vehicle and license-related infractions, (3) assault, battery, and harassment, (4) contempt and court fee debts, (5) drugs, (6) theft, and (7) nuisance and disturbance. For the group closest to the legal limit only one variable—not wearing a seatbelt—shows a statistically significant difference at the 95 percent confidence interval. The same is true for the groups at the higher BAC level, with differences in parking meter violations also showing a difference. Overall, the similarity of the two groups on pre-treatment characteristics buttresses the credibility of the research design strategy resulting in unbiased estimates of the causal effects of the drunk driving sanctions.

---

138. Technically, under the Neyman-Holland-Rubin causal model, the validity of a counterfactual is unverifiable, since it is not possible to observe two outcomes at the same time with the same groups. This issue is an important feature of the "fundamental problems of causal inference," discussed in greater depth in Paul W. Holland, *Statistics and Causal Inference*, 81 J. AM. STAT. ASS'N 945, 947 (1986). The consequence, in this experimental or quasi-experimental framework, is that we can never directly measure the causal effect of the treatment on an outcome, but can only make causal inferences. Thus, verifying if the groups are similar using pre-treatment observables in the data is what can be done to make valid causal inferences about the effect of the treatment on the outcome.

Figure 3: Balance on Pre-Treatment Observable Characteristics (Garland County, Arkansas)

Figure 3(a): Covariate Balance for $0.065 \leq BAC < 0.94$

	CASES	BELOW	ABOVE
Demographic			
Age	576	35.05	35.94
Weight	556	169.63	168.68
Gender	575	297	149
Minorities	576	60	30
Race (Each Race Specified)	576	NA	NA
Traffic-Related			
Careless Driving	576	45	28
Disobey Stop/Yield Line	576	36	20
Disobeyed Traffic Device	576	19	7
Driving Left of Center	576	17	6
Fail to Yield/Give Right of Way	576	9	9
Following Too Closely	576	26	10
Parking Meter Violation	576	20	11
Speeding (Speed Unspecified)	576	66	29
Speeding 0-10 MPH Over Limit	576	32	17
Speeding 11-20 MPH Over Limit	576	29	15
Vehicle/License-Related			
Driving w/ Susp. License	576	43	20
Driving w/ Susp. License - Prev. DWI	576	7	6
Expired Drivers License	576	14	7
Improper/Expired Tags or Reg.	576	35	19
No Drivers License	576	36	18
No Liability Insurance	576	110	58
No Seatbelt	576	65	45
No Vehicle License Plate	576	43	15
Assault/Battery/Harassment			
Battery - 3rd Degree	576	26	14
Domestic Battery - 3rd Degree	576	18	7
Harassment	576	10	7
Terroristic Threatening	576	16	10
Contempt/Fees			
Contempt	576	80	36
Failure to Appear	576	82	35
Public Defender User Fee	576	19	5
Drugs			
Posession of Controlled Substance	576	15	7
Theft			
Shoplifting - 1st Offense	576	13	10
Theft of Property - Misdemeanor	576	15	8
Nuisance/Disturbance			
Criminal Mischief - 1st Offense	576	9	6
Criminal Trespass	576	15	5
Disorderly Conduct	576	27	10
Public Intoxication	576	71	45
TOTAL	576	388	188

P-value

Figure 3(b): Covariate Balance for $0.13 \leq BAC < 0.17$

	CASES	BELOW	ABOVE
Demographic			
Age	209	36.11	35.26
Weight	203	172.14	167.19
Gender	209	80	79
Minorities	209	19	16
Race (Each Race Specified)	209	NA	NA
Traffic-Related			
Careless Driving	209	16	17
Disobey Stop/Yield Line	209	6	11
Disobeyed Traffic Device	209	4	4
Driving Left of Center	209	5	3
Fail to Yield/Give Right of Way	209	4	6
Following Too Closely	209	10	2
Parking Meter Violation	209	3	8
Speeding (Speed Unspecified)	209	17	16
Speeding 0-10 MPH Over Limit	209	9	9
Speeding 11-20 MPH Over Limit	209	4	7
Vehicle/License-Related			
Driving w/ Susp. License	209	16	11
Driving w/ Susp. License - Prev. DWI	209	1	4
Expired Drivers License	209	4	4
Improper/Expired Tags or Reg.	209	12	9
No Drivers License	209	8	12
No Liability Insurance	209	33	33
No Seatbelt	209	15	29
No Vehicle License Plate	209	14	8
Assault/Battery/Harassment			
Battery - 3rd Degree	209	7	8
Domestic Battery - 3rd Degree	209	4	2
Harassment	209	3	5
Terroristic Threatening	209	6	4
Contempt/Fees			
Contempt	209	24	23
Failure to Appear	209	27	25
Public Defender User Fee	209	5	2
Drugs			
Posession of Controlled Substance	209	4	6
Theft			
Shoplifting - 1st Offense	209	4	8
Theft of Property - Misdemeanor	209	2	5
Nuisance/Disturbance			
Criminal Mischief - 1st Offense	209	2	5
Criminal Trespass	209	5	3
Disorderly Conduct	209	11	7
Public Intoxication	209	17	27
TOTAL	209	103	106

P-value axis: 0.0 0.2 0.4 0.6 0.8 1.0

P-value

The dots show the p-values in comparing the group just below versus the group just at or above the legal limit (in Figure 3(a)) or the enhancement level (in Figure 3(b)) on pre-treatment characteristics. Difference-in-means is used for binary variables. The Kalmogorov-Smirnov (K-S) Test is used for continuous variables, and Fisher's Exact Test is used for small samples. Race is coded dichotomously (minority or non-minority) and categorically for each racial group. A Chi-Squared Test is used for the categorical coding of race. Below and above refer to the number of cases (in the case of dichotomous variables) or means (in the case of continuous variables) below and at or above the BAC threshold where a dichotomous variable is equal to one, respectively. Means are presented for continuous variables (age and weight). The dotted line corresponds to a p-value of 0.05, and the striped line shows a p-value of 0.10.

The results are robust across a host of BAC ranges (bandwidths) close to the two thresholds, with at most one to two additional covariates showing imbalance at wider bandwidths. Because one out of every twenty tests is likely by chance to show up statistically significant at the ninety-five percent confidence interval, the covariate balance is strong. In comparing Garland County's recidivism rates and demographics with data at the state level, I have no strong *a priori* reason to believe that similar results will not be obtained with state-level data.

VII.   EMPIRICAL RESULTS

This section explores the causal effect of the sanctions. There are four important dimensions to the results. First, I examine two outcomes of interest: recidivism and subsequent vehicle crashes. Recidivism is operationalized by examining whether or not a first-time offender has at least one subsequent alcohol test with a BAC level greater than or equal to 0.08. The crashes outcome is whether or not a first-time offender gets into a subsequent vehicle accident, irrespective of his or her BAC level. Second, I analyze whether the sanctions have incapacitation or specific deterrence effects. Incapacitation is measured by examining recidivism and subsequent crash rates during the license suspension period, and the specific deterrence outcome captures the same outcomes starting the day after the license suspension period ends. All results in this section examine recidivism within one year. Third, I show RDD results estimate the effectiveness of sanctions within a legal regime, and DID results, which allow for the cross-temporal comparison of the impact of drunk driving laws. Fourth, results are shown for three distinct periods of drunk driving law.

The three distinct legal regimes are in place from 2001 until 2013. Their key sources of variation is the following:

(1) From 2001 until 2005, those at or just above the legal limit received jail time that typically lasted 6–24 hours, fines and court fees, and a license suspension of 120 days, with a restricted permit allowing them to go to school, work, or court. Courts controlled IIDs, but their uptake was relatively low, since those close to the legal limit received a restricted permit, and since some judges did not use them as sanctions.

(2) In 2005, IIDs, which up until then were under the jurisdiction of courts, became part of the jurisdiction of the Office of Driver Control. As a result, their use increased following the legal reform.

(3) In 2009, the legislature voted to have mandatory victim panels for all who were convicted of DWI, and no restricted permits for anyone with a

license suspension. As a result, a person convicted of DWI could only drive legally with an IID.
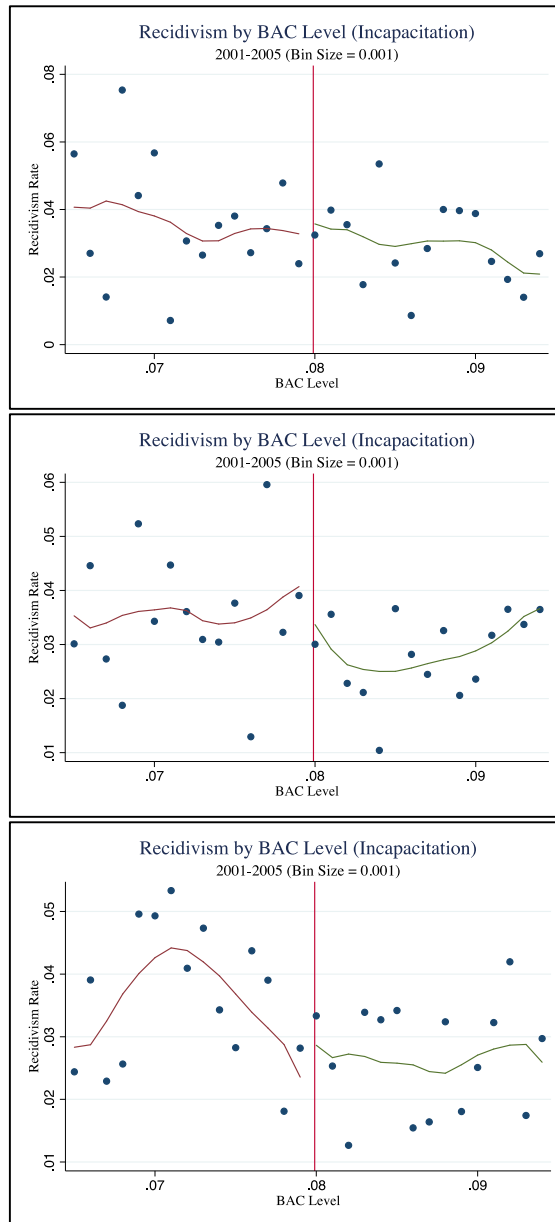
This institutional background sets the backdrop to examine whether the reforms achieved incapacitation and specific deterrence.

A. *Incapacitation Effects*

Do sanctions have an incapacitation effect on drunk driving offenders? I begin by comparing recidivism rates during the license suspension period for drivers with a BAC close to the legal limit of 0.08. Before turning to the estimates, I provide graphical evidence of the relationship between BAC and recidivism. Figure 4 shows mean DWI recidivism rates by BAC level for first-time offenders while their licenses were suspended.[139] Particularly noteworthy is that the difference in recidivism rates for those just below versus those at or just above the legal limit is most apparent during the 2009–2013 period. The nonparametric loess regression lines to the left and right of the 0.08 threshold have the greatest distance from each other during this time period, with lower recidivism rates occurring for those who were just at or above the legal BAC limit.

---

139. The plots are also referred to as the conditional expectation function.

**Figure 4: DWI Recidivism by BAC During
the License Suspension Period at the Legal Limit**

The dots show the mean recidivism rates by BAC level for first-time offenders during the license suspension period. The data is fitted to a loess smoother on either side of the legal BAC limit of 0.08.

The trends in the plots are largely borne out in the difference in means estimates presented in Table 2. The point estimate of -0.9 percentage points, that corresponds to the plot in Figure 4 during 2009–2013 period, has the highest magnitude of any estimate across the three periods when using the 120-day license suspension period, and is statistically significant at conventional levels ($p = 0.04$).[140] The magnitudes of the point estimates and standard errors are slightly lower for the 120 day license suspension period in 2001–2005 and 2005–2009 (-0.7 percentage points for both periods), and their respective confidence intervals slightly are higher than the 2009-2013 period ($p = 0.11$ and $p = 0.10$, respectively). The magnitudes of the point estimates are highest during the 2009–2013 period (-0.9 and -1.2 percentage points using a 120 and 180 day recidivism window, respectively), and the respective p values of 0.04 and 0.03 are both statistically significant at the ninety-five percent confidence intervals. Baseline recidivism rates for the control groups are almost equal for the 120-day and 180-day periods during the 2009–2013 sentencing regime. I examined whether effects might be related to the duration of the incapacitation period, independent of the sanctions.

Table 2: The Incapacitation Effect of DWI Sanctions at the Legal Limit on Recidivism

	2001–2005 120 days	2005–2009 120 days	2009–2013 120 days	2009–2013 180 days
	$0.065 \leq BAC < 0.095$			
Estimate	-0.007	-0.007†	-0.009*	-0.012*
Std. Error	0.004	0.004	0.005	0.005
p	0.11	0.10	0.04	0.03
Baseline Mean	-0.004	-0.003	-0.005	-0.006
Obs.	6750	7586	6255	6042

This table presents the effect of sanctions in reducing recidivism for first-time offenders with BAC levels close to the legal limit during the license suspension period. Estimates report the local average treatment effect using difference in means for the 0.065–0.094 BAC bandwidth. The baseline mean is the recidivism rate for the control group (the group just below the legal limit of 0.08). Although the license suspension during the 2009–2013 period

---

140. The 2009 law legislated an increase in license suspensions from 120 to 180 days for all first-time offenders. I thus present estimates during the 2009–2013 period for 120 and 180 days for all specifications that examine incapacitation effects. All dot-plots show the 120 day suspension period for the 2009–2013 period.
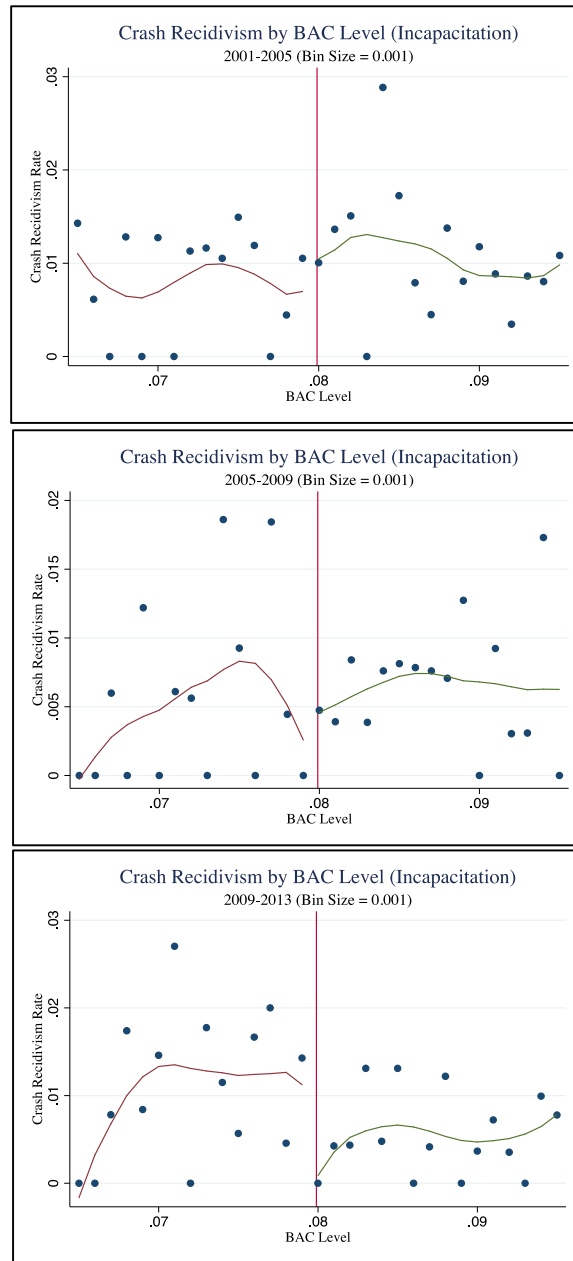
is 180 days, recidivism during 120 days is also reported for comparison with the other periods.
[*]$p \leq 0.05$
[†]$0.05 < p \leq 0.10$

The effects are not as strong when examining whether sanctions have an incapacitation effect on subsequent vehicle crashes by first-time offenders. The plots in Figure 5 show a surprising pattern during 2001–2005. For those sanctioned at the legal limit, while their license is suspended, the probability of a vehicle accident increases, although the estimate is not statistically significant at conventional levels. Based on the plots, the size of this increase appears to diminish during the 2005–2009 period, followed during the 2009–2013 period by a dramatic reduction in crashes for those who are sanctioned.

Figure 5: Subsequent Crashes by BAC During the License Suspension Period

The dots show the mean vehicle crash recidivism rates by BAC level for first-time offenders during the license suspension period. The data is fitted to a loess smoother on either side of the legal BAC limit of 0.08.

Table 3 confirms this pattern with positive point estimates of 0.3 and 0.0 percentage points for the 2001 and 2005 legal reform periods, respectively. Although the estimates are positive, the increased crashes in the first two periods are not statistically significant at conventional levels, and the confidence interval for both estimates easily crosses below zero. The 2009–2013 period is the only period where the sanctions have an incapacitation effect on crashes. The sanctions in place during that period had a strong effect. The crash rate for those sanctioned was 0.5 to 0.6 percentage points lower than those who were not sanctioned. This effect amounts to a fifty-five percent reduction in crashes for the group just above the legal limit relative to those just below.

Table 3: The Incapacitation Effect of DWI Sanctions at the Legal Limit on Subsequent Vehicle Crashes

	2001–2005 120 days	2005–2009 120 days	2009–2013 120 days	2009–2013 180 days
$0.065 \leq BAC < 0.095$				
Estimate	0.003	0.000	-0.005[*]	-0.006[*]
Std. Error	0.002	0.002	0.002	0.003
p	0.27	0.97	0.02	0.02
Baseline Mean	0.001	0.000	-0.003	-0.003
Obs.	6750	7586	6255	6044

This table presents the effect of sanctions in reducing subsequent vehicle crashes for first-time offenders with BAC levels close to the legal limit during the license suspension period. Estimates report the local average treatment effect using difference in means for the 0.065–0.094 BAC bandwidth. The baseline mean is the recidivism rate for the control group (the group just below the legal limit of 0.08). Although the license suspension during the 2009–2013 period is 180 days, recidivism during 120 days is also reported for comparison with the other periods.
[*] $p \leq 0.05$
[†] $0.05 < p \leq 0.10$

Before making cross-temporal comparisons, I examine whether an incapacitation effect also exists at the enhanced BAC level of 0.15. That threshold triggers a license suspension increase from 120 to 180 days, and the comparison is made for recidivism rates for the sixty-day period when the group just below the enhancement level is able to drive without a suspended license, while the group with BACs at or above 0.15 still had a suspended license. Table 4 shows the difference in means in recidivism rates and crashes

for first-time offenders close to the 0.15 BAC level. Because the state legislature passed legislation abandoning the enhancement in 2009, time periods are restricted to before 2009. In 2003, the state legislature took away restricted permits for first-time offenders with BAC levels at or above 0.15,[141] so I include separate estimates for the 2001–2003 and 2003–2005 periods.[142] Although circumscribed to this particular context and this BAC level, the finding contrasts with scholars who have discussed the strongly punitive nature and effectiveness of license suspensions in curbing recidivism.[143] Unfortunately, the relative scarcity of data precludes the possibility of doing a well-informed analysis of the causal effect of the enhancement on subsequent crashes, but perhaps surprisingly, we see that the license suspension *increases* recidivism in the 2001–2003 period by 0.6 percentage points.

Table 4: The Incapacitation Effect of DWI Sanctions at the Enhancement Level on Recidivism

	2001–2003	2003–2005	2005–2009
0.14 ≤ BAC < 0.16			
Estimate	0.006†	0.002	-0.001
Std. Error	0.004	0.004	0.003
p	0.08	0.57	0.78
Baseline Mean	0.003	0.001	0.000
Obs.	3445	3107	5377

This table presents difference in means results for the effect of license suspension sanctions for first-time offenders at the 0.15 BAC threshold on recidivism. Specifications compare recidivism rates for days 151–210 for both groups, when those just below the enhancement level do not have an active license suspension, and those above the enhancement level do have a suspension. During all three periods, first-time offenders with a BAC < 0.15 receive a 120-day license suspension and a restricted permit, allowing them to drive to a few locations, including work and school. From 2001–2003, first-time offenders with BAC ≥ 0.15 received a 180-day license suspension and restricted permit. From 2003–2009, first-time offenders with BAC ≥ 0.15 could not obtain a restricted permit. Their only option to drive legally was to have an IID. In 2005, the Arkansas legislature transferred jurisdiction over IIDs from the courts to the Office of Driver Control.

---

141. Act 1779, H.B. 2716, 84th Gen. Assemb., Reg. Sess. (Ark. 2003), http://www.arkleg.state.ar.us/assembly/2003/R/Acts/Act1779.pdf.

142. I also examine the effect of the enhancement for 180 days for both the treatment and control groups, and find no effect of the enhancement on recidivism. Results are available from the author.

143. *See generally* Nichols & Ross, *supra* note 69; Peck et al., *supra* note 69, at 57; Voas, *supra* note 69.

$^*p \leq 0.05$
$^†0.05 < p \leq 0.1$

The results thus far have focused on the effect of sanctions within a legal regime for drunk driving. I now turn to making comparisons across different legal regimes, examining the effect of legal reforms on incapacitation. Table 5 shows DID estimates that give the causal effect of the 2005 and 2009 reforms to the drunk driving statute on recidivism and crashes. Although the RDD point estimates were suggestive of IIDs having an effect across time periods, the confidence interval for the DID estimates overlaps with zero for the reforms in 2009 ($p = 0.62$ and $p = 0.30$, respectively). This result undermines support for the hypothesis that the 2009 reforms—which included requiring an IID to drive legally during the license suspension period and mandatory attendance at victim panels—had an effect on reducing recidivism relative to the 2005 period.

Table 5: The Incapacitation Effect of the 2005 and 2009 Drunk Driving Reforms on Recidivism and Vehicle Crashes (Difference-in-Differences Results)

	Recidivism			Subsequent Crashes		
	2005 120 days	**2009 120 days**	**2009 180 days**	**2005 120 days**	**2009 120 days**	**2009 180 days**
	0.065 ≤ BAC < 0.095					
Estimate	-0.002	-0.003	-0.007	-0.002	0.006†	0.007*
Std. Error	0.006	0.006	0.007	0.003	0.003	0.004
p	0.72	0.62	0.30	0.46	0.06	0.06
Obs.	13028	13115	13118	13028	13118	13102

This table presents difference-in-differences (DID) results for the incapacitation effect of the 2005 and 2009 legal reforms on recidivism and crashes. All specifications include year-county fixed effects using robust clustered standard errors, clustered by year and county. Recidivism for incapacitation specifications is equal to one if a driver has a subsequent drunk driving offense (BAC $\geq$ 0.08) at any time during the license suspension period. Subsequent crashes are equal to one if a driver was stopped for an alcohol test and had a subsequent crash during the license suspension period. Although the license suspension during the 2009–2013 period is 180 days, recidivism during 120 days is also reported for comparison with the other periods.
$^*p \leq 0.05$
$^†0.05 < p \leq 0.10$

The 2009 sanctions, however, do have a *positive* effect on subsequent crashes. The estimates in the fifth and sixth columns show an increase in vehicle crashes ranging from 0.6 to 0.7 percentage points, depending on whether one chooses 120 days or 180 days, respectively, as the duration for a first-time offender to have a subsequent crash in the 2009 period. The result might be suggestive of increased nervousness or something inherent to the machine that could create a difficult situation when driving. The point estimates are borderline statistically significant at the ninety percent and ninety-five percent confidence levels, respectively. The result lends support that an aspect of the reform resulted in fewer vehicle accidents. Evidence for this mechanism is described in greater detail in Section VII(C).
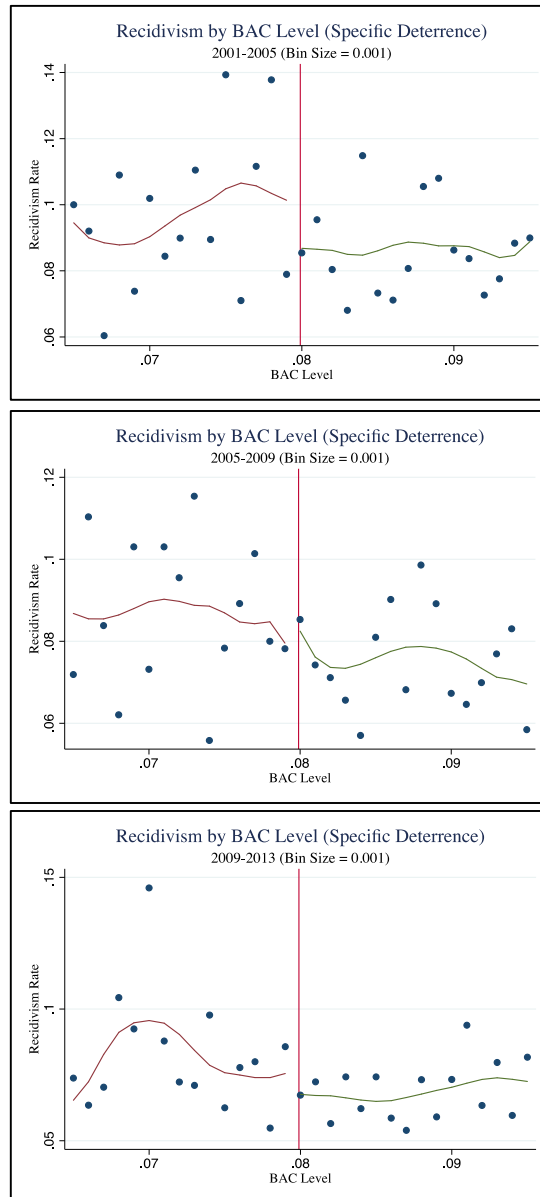
B. *Specific Deterrence Effects*

The results presented for specific deterrence estimate recidivism and crashes within a year of the license suspension being lifted. I also examine the same outcomes for a period ranging from one month up to three years, to examine the robustness of the results.[144]

The plots in Figure 6 show the means of one year DWI recidivism rates for first-time offender BAC levels close to the 0.08 threshold. The presence of a discontinuity in some cases appear easily visible, which is suggestive of the sanctions having specific deterrence effects.

---

144. Results are available from the author upon request. The upper bound of three years is chosen because of censoring of observations at the end of the data set for the 2009–2013 period.

Figure 6: Subsequent Crashes by BAC After the License Suspension Period at the Legal Limit (Specific Deterrence)

The dots show the mean one-year recidivism rates by BAC level for first-time offenders starting the day the license suspension is lifted. The data is fitted to a loess smoother on either side of the legal BAC limit of 0.08.

Estimates in Table 6 are consistent with this trend in the graphical data. All of the results, irrespective of the legal regime, show an effect on specific deterrence ranging from 1.3 to 1.4 percentage points. This variation contrasts with policy reforms with varied forms of incapacitation that have created contrasting results on recidivism and vehicle crashes.

Table 6: The Specific Deterrence Effect of DWI Sanctions at the Legal Limit on Recidivism

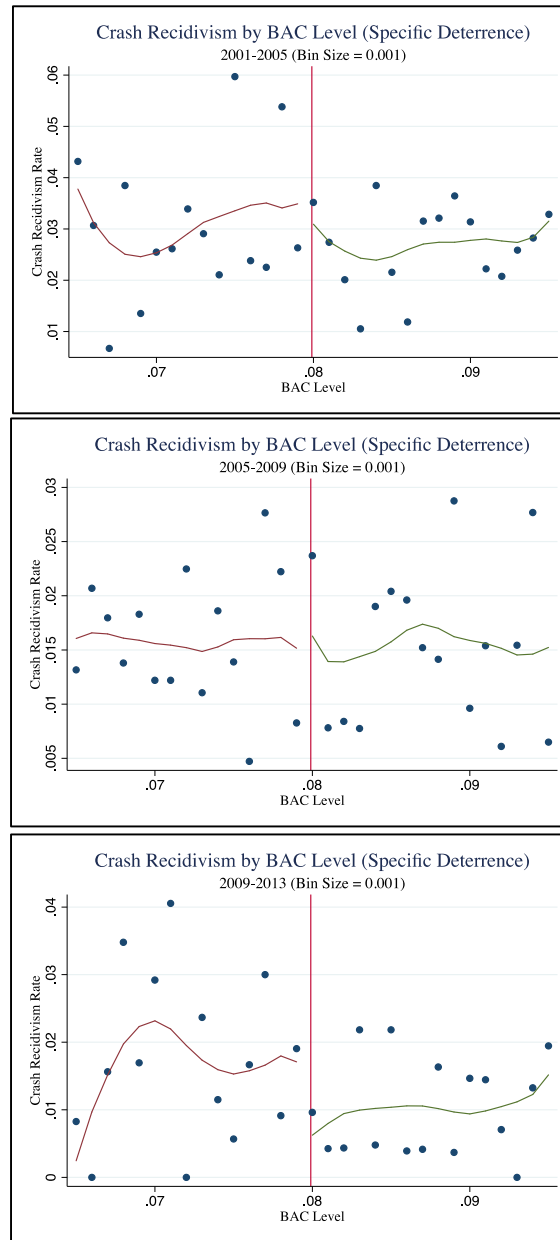|	**2001–2005**	**2005–2009**	**2009–2013**
$0.065 \leq BAC < 0.095$			
Estimate	-0.013[†]	-0.013[*]	-0.014[†]
Std. Error	0.007	0.006	0.007
p	0.07	0.04	0.06
Baseline Mean	-0.006	-0.006	-0.007
Obs.	6750	7586	6036

This table presents the specific deterrence effect of sanctions in reducing recidivism for first-time offenders with BAC levels close to the legal limit for one year after a license suspension is lifted. Estimates report the local average treatment effect using difference in means for the 0.065–0.094 BAC bandwidth. The baseline mean is the recidivism rate for the control group (the group just below the legal limit of 0.08).
[*] $p \leq 0.05$
[†] $0.05 < p \leq 0.10$

Figure 7 shows the same specific deterrence effect of the sanctions, but with the outcome of subsequent crashes. The pattern is similar to the results for the incapacitation effect of crashes—a discontinuity is visible only in the 2009–2013 period, with subsequent crashes being lower for those who experienced sanctions. The trend is suggestive of the 2009 reforms not only having a greater impact on crashes, but also having a longer lasting effect than the previous reform period.

**Figure 7: Subsequent Crashes by BAC After the License Suspension
Period at the Legal Limit (Specific Deterrence)**

The dots show the mean one-year crash recidivism rates by BAC level for first-time offenders starting the day the license suspension is lifted. The data is fitted to a loess smoother on either side of the legal BAC limit of 0.08.

The estimates in Table 7 confirm the trends in the plots in Figure 7. In examining whether there is a causal effect of the sanctions on subsequent crashes, we only see an effect in the 2009–2013 period. Specifically, I find a statistically significant reduction ($p = 0.03$) in crashes of 0.7 percentage points. This accounts for a forty-four percent decrease in crashes for those within the 0.065 to 0.095 BAC range.

Table 7: The Specific Deterrence Effect of DWI Sanctions at the Legal Limit on Subsequent Crashes

	2001–2005	2005–2009	2009–2013
0.065 ≤ BAC < 0.095			
Estimate	-0.004[†]	-0.002	-0.007[†]
Std. Error	0.004	0.003	0.003
p	0.38	0.41	0.03
Baseline Mean	-0.002	-0.001	-0.004
Obs.	6750	7586	6044

This table presents the specific deterrence effect of sanctions in reducing subsequent crashes for first-time offenders with BAC levels close to the legal limit for 1 year after a license suspension is lifted. Estimates report the local average treatment effect using difference in means for the 0.065–0.094 BAC bandwidth. The baseline mean is the recidivism rate for the control group (the group just below the legal limit of 0.08).
[*] $p \leq 0.05$
[†] $0.05 < p \leq 0.10$

Table 8 shows the results of the effect of a two-month license suspension enhancement on recidivism and crashes. All but one of the specifications have confidence intervals that cross zero with p values ranging from 0.32 to 0.98. Consequently, only one of the estimates reaches the range of statistical significance at conventional levels. The confidence intervals are not significantly tight to rule out a null finding of no effect, or a "tight zero."[145] The two month enhancement thus appears to be a relatively weak deterrent for this group of first-time offenders stopped for their first offense at nearly twice the legal limit. Other possibilities could be that the severity of other aspects of the sanctions has reached a level where marginal differences in severity are unlikely to have significant effects. The result raises the

---

145. To check that the lack of an effect for the 2001–2003 and 2003–2005 periods are not driven by a lack of statistical power, I also run pooled results from 2001–2005 for recidivism and crashes, and find no statistically significant effect.

challenge of regulating "crimes of addiction" where classical models of deterrence may not be effective in reducing recidivism.

Table 8: Incapacitation vs. Specific Deterrence of DWI Sanctions at the Enhancement Level on Recidivism and Subsequent Vehicle Crashes

	Recidivism			Vehicle Crashes		
	2001–2003	2003–2005	2005–2009	2001–2003	2003–2005	2005–2009
$0.14 \leq BAC < 0.16$						
Estimate	0.006[†]	0.002	0.000	0.002	0.001	0.001
Std. Error	0.003	0.004	0.002	0.002	0.001	0.001
p	0.08	0.58	0.93	0.38	0.32	0.46
Baseline Mean	0.003	0.001	0.000	0.001	0.001	0.000
Obs.	3696	2731	6585	3793	2791	6761

This table presents difference in means results for the specific deterrence effect of sanctions for first-time offenders at the 0.15 BAC threshold on recidivism. Specifications compare recidivism rates for days 151–210 for both groups, when those just below the enhancement level do not have an active license suspension, and those above the enhancement level do have a suspension. During all three periods, first-time offenders with a BAC < 0.15 receive a 120-day license suspension and a restricted permit, allowing them to drive to a few locations, including work and school. From 2001–2003, first-time offenders with BAC ≥ 0.15 received a 180-day license suspension and restricted permit. From 2003–2009, first-time offenders with BAC ≥ 0.15 could not obtain a restricted permit. Their only option to drive legally was to have an IID. In 2005, the Arkansas legislature transferred jurisdiction over IIDs from the courts to the Office of Driver Control.
[*] $p \leq 0.05$
[†] $0.05 < p \leq 0.10$

Comparing the effects of the sanctions across the legal regimes, I find that the DID results that provide estimates of the causal effect of the 2005 and 2009 laws by and large do not reach levels of conventional statistical significance for any specification, irrespective of the outcome. However, worth noting is that one point estimate from the 2001–2003 period, in comparison with the 2003–2005 period, did have a 0.6 percentage points on recidivism. Nevertheless, this conclusion should be taken with some reservation.

Table 9: The Specific Deterrence Effect of the 2005 and 2009 Drunk Driving Reforms on Recidivism and Vehicle Crashes (Difference-in-Differences Results)

	Recidivism		Vehicle Crashes	
	2005	**2009**	**2005**	**2009**
	$0.14 \leq BAC < 0.16$			
Estimate	0.006	0.013	-0.007	-0.002
Std. Error	0.016	0.011	0.008	0.005
p	0.69	0.24	0.36	0.64
Obs.	8231	10372	8231	10372

This table presents difference-in-differences (DID) results for the specific deterrence effect of the 2005 and 2009 legal reforms on recidivism and crashes. All specifications include year-county fixed effects using robust clustered standard errors, clustered by year and county. Recidivism for incapacitation specifications is equal to one if a driver has a subsequent drunk driving offense (BAC $\geq$ 0.08) at any time during the license suspension period. Subsequent crashes are equal to one if a driver was stopped for an alcohol test and had a subsequent crash during the license suspension period. Although the license suspension during the 2009–2013 period is 180 days, recidivism during 120 days is also reported for comparison with the other periods.
[*] $p \leq 0.05$
[†] $0.05 < p \leq 0.10$

Table 10 summarizes the key results of the analysis in tabular form. A number of findings emerge from the overall results.

First, Arkansas drunk driving sanctions—under certain conditions—can have both specific deterrence and incapacitation effects, however because the deterrence effect has been relatively constant over time, it might be the case that focusing on the efficacy of incapacitation might provide a greater effect in reducing subsequent crime and crashes.

Second, DID results, which permit the comparison of the legal regimes across time, offer promise for the efficacy of non-carceral sanctions in terms of reducing crashes. The non-carceral reforms implemented by the law, which included stronger incentives to use IIDs and mandatory victims panels, were certainly effective in reducing crashes while a first-time offender's license suspension is in effect. Their impact on recidivism is mixed.

Additionally, the effectiveness of license suspensions is mixed. A license suspension enhancement of sixty extra days at the 0.15 BAC threshold has limited incapacitation or deterrent effect on recidivism or crashes. A sixty-day increase in license suspension was part of the 2009–2013 reforms. Because I cannot cleanly isolate the effectiveness of the license suspension

change at the legal limit, I cannot rule out that the license suspension may have contributed to decreases in recidivism during the license suspension period and to crash reductions in the year after the suspension.

Furthermore, recidivism reductions through incapacitation during the 2001 legal regime are likely the consequence of a reduction in the legal limit for drunk driving, rather than the efficacy of the punishments during that time.

Finally, the relative absence of statistically significant RDD results for incapacitation and deterrence during the 2001 and 2005 legal regimes is surprising. First-time DWI offenders are given at least six to twenty-four hours of jail, fines, and a license suspension. Their probability of reoffending or getting into a subsequent vehicle crash is not statistically distinguishable from the probability of those who received an alcohol test but no sanctions. This result is suggestive of a number of possibilities, the two most likely of which include (1) the group whose BAC was just below may have been "scared straight" by the interaction with law enforcement and the experience of the alcohol testing, attenuating the relative local average treatment effects of the sanctions for the comparison group at or just above legal limit; or (2) the inefficacy of the sanctions. Adjudicating between these two mechanisms is an important line of inquiry to pursue in future research.

Table 10: Summary of Results that Achieve Statistical Significance

Incapacitation (Recidivism & Crashes During License Suspension)					
Specification	Outcome	2001–2005 (120 days)	2005–2009 (120 days)	2009–2013 (120 days)	2009–2013 (180 days)
RDD	Recidivism	No	Yes†	Yes*	Yes*
RDD	Crashes	No	No	Yes*	Yes*
DID	Recidivism	✕	No	No	No
DID	Crashes	✕	No	Yes†	Yes*
Specific Deterrence (Recidivism & Crashes After License Suspension)					
		2001–2005	2005–2009	2009–2013	
RDD	Recidivism	Yes†	Yes*	Yes†	
RDD	Crashes	Yes†	No	Yes†	
DID	Recidivism	✕	No	No	
DID	Crashes	✕	No	No	

RDD indicates a regression discontinuity design specification that compares a group just below the legal limit for drunk driving with those at or just above the threshold within one of the three legal regimes. RDD specifications are for difference in means results. DID indicates a difference-in-differences specification, which include year-county fixed effects using robust clustered standard errors, clustered by year and county. Recidivism for incapacitation specifications is equal to one if a driver has a subsequent drunk driving offense (BAC $\geq$ 0.08) at any time during the license suspension period. For incapacitation, crashes are equal to one if a driver was stopped for an alcohol test and had a subsequent crash during the license suspension period. Because the legislature increased the duration of the license suspension during the 2009–2013 period to 180 days, I also include a specification of comparable duration (120 days) with the two previous periods. Recidivism and crashes for the specific deterrence specifications indicate whether an individual reoffended between one day and one year after the end of the license suspension period. Shaded gray cells with "Yes" indicate the specification had an effect at conventional levels of statistical significance ($p \leq$ 0.1). All specifications are for the bandwidth where: $0.065 \leq BAC < 0.095$. No specifications at the enhancement level had effects at conventional levels of statistical significance.
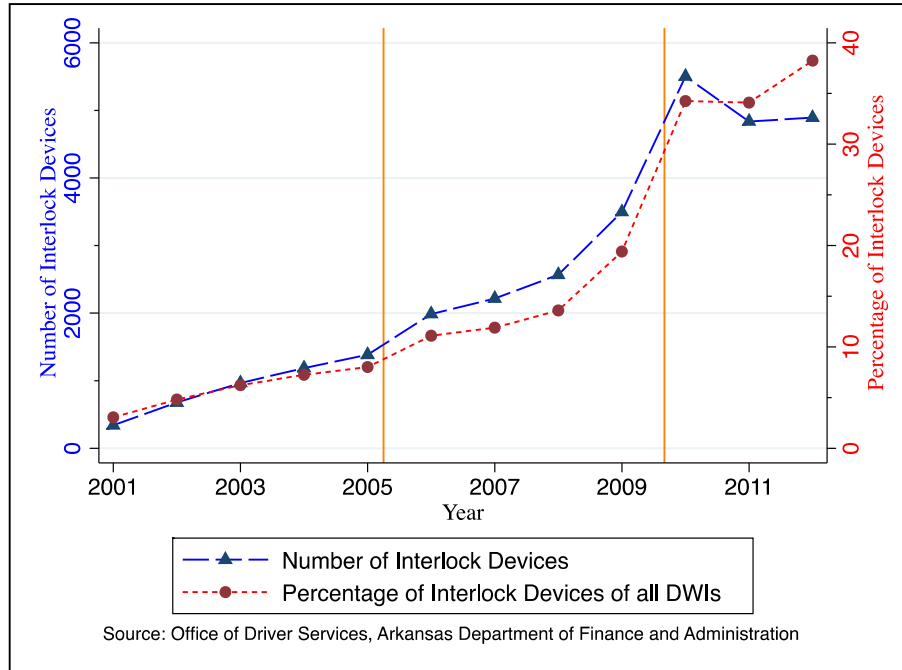* $p \leq 0.05$
† $0.05 < p \leq 0.10$

C. Evidence for a Causal Mechanism

What drives these results? One of the main changes in law that took place in 2009 occurred was the incentive to install an IID, now required for anyone convicted of drunk driving to drive legally. To analyze the effectiveness of the device I turn to some descriptive evidence of the spread of IIDs in Arkansas, in order to discuss the mechanism that could be driving the declines in recidivism and crashes resulting from the 2009 reform. Figure 8 shows IID adoption in Arkansas from 2001 to 2012. The largest annual increase in IID use took place between 2009 and 2010 (from 3,497 IIDs in 2009 to 5,502 in 2010), almost certainly as a consequence of the 2009 legislation that outlawed restricted permits without the devices. For the first time, drivers just above the legal limit (with BAC levels below 0.15) had to install the devices in order to drive legally. Unfortunately, the Office of Driver Control would not release individual-level data on IIDs that would shed light on the mechanisms that are achieving effects on recidivism and crashes through incapacitation, but also through deterrence.

**Figure 8: Ignition Interlock Adoption in Arkansas
2001–2012**

Source: Office of Driver Services, Arkansas Department of Finance and Administration

Note: Vertical lines show the timing of the passage of Act 1234, which the legislature approved on March 24, 2005, and Acts 946 and 1293, which took effect on August 1, 2009, respectively. Act 1234 amended section 5-65-118 of the Arkansas Code, and took away jurisdiction over ignition interlock devices from the courts and gave it to The Office of Driver Services (a division of the Arkansas Department of Finance and Administration). On August 1, 2009, Acts 946 and 1293 took away restricted permits, which gave limited privileges to those with BAC levels above the legal limit, but below 0.15. As a consequence, the only way for those with a drunk driving conviction to drive legally during the suspension period was to have an ignition interlock device.

The lack of individual-level data opens the possibility that at least three non-mutually exclusive mechanisms may be at play.[146] First, because only some forty percent of DWI offenders have IIDs in their car, one possibility is that some channel related to the devices themselves or the behaviors that frequently accompany the devices is effective in reducing recidivism and

---

146. As a result of lacking individual-level IID data, we are forced to make ecological inferences about the effects of and mechanisms through which the devices may reduce recidivism and subsequent vehicle crashes in Arkansas. The five mechanisms described are thus inclusive of typical issues that arise with making ecological inferences.

subsequent crashes. Some possibilities explaining how this direct mechanism of the IID might have an effect on an individual's behavior might include: (1) *a salience mechanism*, where having the device in a visible location in the car induces the driver to drive more carefully and avoid crashes more than he or she might otherwise, or triggers memories which might result in safer driving and less recidivism; (2) *a signaling device*, to law enforcement officials and others that make detection of illegal driving simpler, ultimately leading law enforcement to detect those driving illegally more easily; (3) a *coercive channel* of the device itself, where breathalyzing into and using the IID results in lower recidivism and safer driving, either by "scaring the person straight," or by forcing the person to breathalyze regularly while driving, or through some other similar means. A second possibility is that those who do not have an IID are driving illegally on a suspended driver's license, and are aware of the heightened consequences not only of driving on a suspended license, but also of driving without an IID. Finally, some unobservable factor or contemporaneous changes that happened around the same time could have driven the result. While I cannot definitively exclude this possibility, in-depth interviews, in combination with checking the stability of observables over time, reveal the low likelihood that these factors are at play. With the exception of this last possibility, for the policymaker, one important implication is that although the mechanism may not be fully understood, increasing the use and prominence of IIDs is likely to be effective in reducing crime and vehicle crashes.

One other noteworthy item from Figure 8 is that Arkansas also experienced a noteworthy increase in the number of interlock devices in 2005. At that time, courts were not using the devices that very much in sentencing, so the legislature shifted jurisdiction of IIDs to the Office of Driver Control. We see from the results that the change itself did not result in statistically significant reductions in recidivism and crashes in the first three years. However, the upward trend in IIDs evident in Figure 8 may have laid the groundwork for the 2009 legislation to be effective**.**

Some scholars have theorized that the best way to administer punishment is in a manner that is swift and certain.[147] These scholars have suggested that administrative agencies are likely to be better in adjudicating and administering punishment in this manner. Yet, at least in this specific case, the reform alone induced by the law does not appear to be enough to have had an effect on reducing recidivism and crashes. Only with the combination of changing the underlying structure of behavioral incentives did the

---

147. KLEIMAN, *supra* note 125, at 91.

administrative policy likely (although one cannot say definitively) start to have effects in reducing recidivism and crashes.

VIII.   POLICY AND LEGAL REFORM IMPLICATIONS

A. *Drunk Driving Law*

Taken together, the results reveal a number of patterns that are suggestive of legal and policy reform opportunities. First, IIDs are effective in reducing vehicle accidents. However, they exert more of an "incapacitation" effect than working through specific deterrence. IIDs provide a low-cost and effective way to reduce recidivism and subsequent vehicle crashes. Although the precise mechanisms through which they are achieving effects is an important line of inquiry for future research, the implication for the policymaker is that IIDs—whether through a direct or indirect channel—have strong incapacitating effects in terms of recidivism and subsequent crashes for first-time offenders. Under certain conditions, they also have some specific deterrence effects as well.

Second, punishments at higher BAC levels in terms of enhanced license suspensions appear not to be effective in reducing recidivism or subsequent vehicle crashes. While I cannot precisely identify the mechanism as to why the license suspensions may not be effective, the evidence is suggestive that the group targeted with the enhancement (whose BAC is almost twice or more than twice the legal limit) might reduce their subsequent offenses and crashes with "harder" forms of non-carceral incapacitation, like IIDs.

Third, most penalties at the legal limit appear to be reducing recidivism and subsequent crashes through a specific deterrence, rather than an incapacitation mechanism. The strongest form of incapacitation appears to be IIDs, and the possibility of using them more, through making them mandatory and establishing a fund for the indigent with the fees that are paid for them by those who are not indigent, would likely increase their effectiveness, ultimately resulting in a benefit for society, since it would likely result in fewer vehicle accidents. Though license suspensions are ineffective at higher levels, low enforcement of license suspensions might be part of the issue of their inefficacy. In addition to targeting harder sanctions toward sub-groups with higher BACs, the possibility of increasing randomized sobriety checks may be a way to increase their efficacy. This policy, however, does not come without costs, since civil liberties issues are likely to come into play.

B. Other Crimes

Beyond drunk driving, non-carceral sanctions are playing an important role in a number of different domains of crime. Yet, without rigorous testing, a healthy skepticism of whether results will hold up is in order. The domain that is likely to be closest to drunk driving (DWI) is "driving under the influence" (in Arkansas, DUI), where the offender is found to have been on illegal drugs while driving. The sanctioning is similar, although at least in Arkansas, sentencing appears to have a bit more discretion. Prosecutors are able to engage in charge bargaining, can drop charges, and judges have discretion because the offenses are eligible for expungement.

Nevertheless, even if there is discretion, if it is well-understood, one can do a variant of the regression discontinuity design conducted in this paper—a fuzzy RDD. The fuzzy RDD could be done where actors in the system adhere to a formula for their actions, and treatment assignment has some probability of taking place, rather than being deterministic as it is in this paper, where the actor will receive the treatment (in this case punishment or a harsher punishment) automatically. While I have not investigated in depth other domains where the regression discontinuity can be deployed, misdemeanor drug offenses more generally seem to be an area that has somewhat formulaic sentencing and where the law creates discontinuities that could possibly be exploited for quasi-experimental designs. Substantively, similar issues of regulating so-called "crimes of addiction" are also an important element of the sentencing regime, where responsiveness to sanctions might be more difficult to obtain for those suffering from drug addiction.

One important point is that the RDD is just one form of quasi-experimental design that can be used for this type of analysis. Other sanctions that are likely to have discontinuities include those mentioned on the incapacitation continuum. Electronic monitoring and probation, in particular, seem to be punishments that are likely to have thresholds in place that determine the time someone is subjected to the sanction. If discontinuities do not exist, one other possibility—which may seem far-fetched at first—is to have judges, probation officers, and others in the criminal justice system who administer sanctions either create discontinuities that would allow for the possibility of testing various thresholds or randomly assign sanctions when they are at the margin or within a range for some offenses where the authority found either sanction to be a possibility for the offender. Courts, agencies, and other authorities have an important legitimating function in the administration of punishment, and there is no doubt that experimental sanctioning could undermine that function. Nevertheless, in settings where the sanctions regime permits options, this type of experimentation might be less problematic from

an ethical standpoint. Finally, natural experiments offer another opportunity to test the effects of non-carceral sanctions in a rigorous manner. Scholars have already used the random assignment of judges to cases to examine the punitive nature of judges.[148] A natural extension of this work would be to test the extent to which different forms of non-carceral sanctioning judges are using affect a range of outcomes, including recidivism.[149]

C. *Beyond Criminal Law*

The intersection of criminal law, criminal procedure, and administrative law is an area where institutional design, administration, and procedural concerns drive varied outcomes in sanctioning. In Sections V and VII, the article discussed the effects of a reform designed to give an agency power over the administration of license suspensions and IIDs. The 2005 transfer of jurisdiction to the Office of Driver Control likely created important pathways for the increased usage of IIDs. As a descriptive matter, the transition speaks to important institutional design and policy implementation questions raised by scholars such as Mark Kleiman and Jerry Mashaw in their respective path-breaking works, *When Brute Force Fails* and *Bureaucratic Justice*.[150] Kleiman discusses how punishment is most effective when it is applied in a swift, non-discriminatory, and severe manner, with severity being the least important of the three. Meanwhile, Mashaw, in the context of social security claims, analyzes the important role that courts versus agencies play in administering sanctions. Courts in some ways can be the antithesis to Kleiman's notion of effective punishment. They tend to be slower than agencies, and in most cases, they have more discretion, leaving the opportunity for greater punishment disparity to occur. While procedural fairness concerns should not be overlooked and should be made central to most questions of policy implementation, perhaps it is not surprising that the agency was more effective in disseminating IIDs more evenly. However, it

---

148. *See, e.g.*, Anna Aizer & Joseph J. Doyle Jr., *Juvenile Incarceration, Human Capital and Future Crime: Evidence from Randomly-Assigned Judges*, 130 Q.J. ECON. 759, 759 (2015); Donald P. Green & Daniel Winik, *Using Random Judge Assignments to Estimate the Effects of Incarceration and Probation on Recidivism Among Drug Offenders*, 48 CRIMINOLOGY 357, 357 (2010).

149. In current research I am conducting with Ryan Sakoda, we are exploiting random assignment of judges, to examine whether judges expunge at different rates, and what the determinants might be of their expungement practices. Conditional on there being variation in their expungement, we can then use this variation to examine the effects of expungement on employment and recidivism. Expungement may be conceived as the "undoing" of a sanction, and thus, could be viewed as a form of "undoing" the severity of a sanction.

150. KLEIMAN, *supra* note 125, at 122; JERRY L. MASHAW, BUREAUCRATIC JUSTICE: MANAGING SOCIAL SECURITY DISABILITY CLAIMS 77 (1983).

was not until the combination of a strong behavioral incentive—the 2009 change in law that mandated the device for all DWI offenders to drive legally—that major reductions in recidivism and vehicle crashes started to take place. The law spurs an important agenda about the conditions under which punishment is effective when delegated to agencies. Although additional testing is needed, it at least raises the possibility that other states might consider giving jurisdiction of license suspensions and IIDs to the agency that controls motor vehicles.

IX.   CONCLUSION

The study is part of a larger agenda oriented toward studying the effectiveness of non-carceral sanctions and varied forms of incapacitation. Understanding the conditions under which they are effective will shed light on the means through which more efficient reallocation of resources can be used for corrections policy. The sanctions also force us to rethink how we conceptualize and measure incapacitation, offering a broader and more continuous notion of the concept, while also allowing for reflection on what the individual is being incapacitated from. The findings also have implications for creating environments, when possible, to do policy evaluation with quasi-experimental methods. This ultimately allows us to make informed policy decisions that are of great consequence to society.

Although we are starting to experience a slow reversal of mass incarceration trends that started in the 1970s, policy positions remain polarized on whether the prison boom has been beneficial for society because it helped reduce crime, or whether it laid the groundwork for higher recidivism and the rupturing of communities. Non-carceral sanctions might offer common ground in this debate, because the offender experiences a sanction, and aspects of the sanctions might benefit society and also help the defendant reintegrate with society. The first step to assessing their effectiveness is to embark on a course of rigorous evaluation of their effects. This article attempts to do so with the case of drunk driving, and finds varied efficacy of different types of non-carceral sanctions. It paves the way for a more rigorous, nuanced, and systematic approach to our criminal justice policy that will hopefully lead to a more informed and beneficial sanctions regime for offenders and society at large.

APPENDIX I: INSTITUTIONAL CONTEXT AND ADDITIONAL DATA DETAIL

A. Measurement Error

Table 11 shows the estimated BAC levels for a given weight and drink combination. The increase in BAC across the scale for a given quantity of alcohol, along with the measurement error of portable breathalyzers increase the credibility of the estimates given in the study.

Table 11: Estimating Blood Alcohol Level
(Based on Weight and Sex)

Table 11(a): Males[151]

Weight	1 drink	2 drinks	3 drinks	4 drinks	5 drinks	6 drinks	7 drinks	8 drinks	9 drinks	10 drinks
100 lbs	0 .043	0 .087	0 .130	0 .174	0 .217	0 .261	0 .304	0 .348	0 .391	0 .435
125 lbs	0 .034	0 .069	0 .103	0 .139	0 .173	0 .209	0 .242	0 .278	0 .312	0 .346
150 lbs	0 .029	0 .058	0 .087	0 .116	0 .145	0 .174	0 .203	0 .232	0 .261	0 .290
175 lbs	0 .025	0 .050	0 .075	0 .100	0 .125	0 .150	0 .175	0 .200	0 .225	0 .250
200 lbs	0 .022	0 .043	0 .065	0 .087	0 .108	0 .130	0 .152	0 .174	0 .195	0 .217
225 lbs	0 .019	0 .039	0 .058	0 .078	0 .097	0 .117	0 .136	0 .156	0 .175	0 .195
250 lbs	0 .017	0 .035	0 .052	0 .070	0 .087	0 .105	0 .122	0 .139	0 .156	0 .173

---

151. *The ABCs of BAC,* NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., http://www.nhtsa.gov/links/sid/ABCsBACWeb/page2.htm (last visited Dec. 19, 2015). One drink is roughly equivalent to one 12-ounce beer, one 4-ounce glass of wine, or 1.5 ounces of distilled spirits. Nat'l Highway Traffic Safety Admin., *Alcohol Screening and Brief Intervention in the Medical Setting*, DOT HS 809 467 (2002), http://www.nhtsa.gov/people/injury/ems/Archive/alcohol_screening/PDFs/Overview.pdf.

Table 11(b): Females

Weight	1 drink	2 drinks	3 drinks	4 drinks	5 drinks	6 drinks	7 drinks	8 drinks	9 drinks	10 drinks
100 lbs	0 .050	0 .101	0 .152	0 .203	0 .253	0 .304	0 .355	0 .406	0 .456	0 .507
125 lbs	0 .040	0 .080	0 .120	0 .162	0 .202	0 .244	0 .282	0 .324	0 .364	0 .404
150 lbs	0 .034	0 .068	0 .101	0 .135	0 .169	0 .203	0 .237	0 .271	0 .304	0 .338
175 lbs	0 .029	0 .058	0 .087	0 .117	0 .146	0 .175	0 .204	0 .233	0 .262	0 .292
200 lbs	0 .026	0 .050	0 .076	0 .101	0 .126	0 .152	0 .177	0 .203	0 .227	0 .253
225 lbs	0 .022	0 .045	0 .068	0 .091	0 .113	0 .136	0 .159	0 .182	0 .204	0 .227
250 lbs	0 .020	0 .041	0 .061	0 .082	0 .101	0 .122	0 .142	0 .162	0 .182	0 .202

Table 11(c): Time Factor Table

Hours since first drink	1	2	3	4	5	6
Subtract from blood alcohol level	0.015	0.030	0.045	0.060	0.075	0.090

B. *Description of the Statistical Models and Estimation Strategy*

1. The Regression Discontinuity Design

More formally, the effect of a treatment—in this case either sanctions at the legal limit for drunk driving or an enhanced sentence at the higher BAC threshold—is estimated by the following reduced form equation:

$$Y_{it} = \alpha + \beta F_{it} + \gamma T_{it} + u_i.$$

I am interested in two outcomes of interest: recidivism and subsequent vehicle crashes. In the case of recidivism at the legal limit, the outcome of interest, Y_{it}, is equal to one when an individual i receives at least a second

alcohol test with a BAC result of 0.08 or higher within some time period, t. F_{it} is the individual's BAC level for their first alcohol test, given at some time, t, T_{it} is a binary treatment indicator, where:

$T_{it} = 1$ if $F_{it} \geq 0.08$
$T_{it} = 0$ if $F_{it} < 0.08$,

and u_i is a disturbance term. The parameter of substantive interest is γ, the slope of which indicates the direction and magnitude of the treatment effect, which is the causal effect of sanctions at the legal limit on recidivism. The model is the same for estimating vehicle crashes, except the outcome, Y_{it}, is whether an individual who has had an official breath test at least once, has a subsequent breath test and is the driver in a subsequent vehicle accident, within some time period, t. I recently obtained the entire database of reported vehicle crashes in Arkansas, and plan to use this data to examine crashes that occur in the absence of an official alcohol test. The reported effects of the sanctions on subsequent crashes are thus likely to be conservative. The equation is also the same for the enhanced license suspension, except the treatment indicator, T_{it}, is equal to 1 if $F_{it} \geq 0.15$, or 0 if $F_{it} < 0.15$.

In order to estimate treatment effects, I rely on three core specifications: (1) difference-in-means; (2) local linear regression; and (3) regressions with polynomials in order to determine the robustness of the results across these specifications. The local linear regression is estimated with the following equation:

$$Y_{it} = \alpha + \beta F_{it} + \gamma T_{it} + \delta F_{it} T_{it} + u_i.$$

The terms are the same as the reduced form, except the equation also includes $\delta F_{it} T_{it}$, an interaction term between F_i and the treatment indicator, where the treatment effect is the difference in slopes of an ordinary least squares regression of the points to the left and those to the right of the threshold. Regressions with polynomials offer a flexible form designed to fit the entire range of the data, in which instead of an interaction term, T_{it}^2, T_{it}^3 T_{it}^4, T_{it}^5, etc. terms are added to the specification, depending on the order of the polynomial.[152] Local linear and polynomial specifications all use clustered-robust standard errors, clustered at the 0.01 BAC level, since that is the level at which sentencing is determined. These standard errors thus account for heteroskedasticity and autocorrelation. Although estimation

---

152. For a more extensive treatment of estimation strategies with RDD, see generally Imbens & Lemieux, *supra* note 91; Lee & Lemieux, *supra* note 90; van der Klaauw, *supra* note 91.

techniques are still an emerging area for RDD, scholars at this time largely concur that showing the results of multiple specifications is a good practice.[153]

2. Difference-in-Differences Estimation

The conventional DID estimator relies on a linear parametric model, that typically compares changes in laws over different time periods. This model compares RDDs in multiple time periods, and some have referred to this design as the "difference in discontinuities" (or the "diff-in-disc") design.[154] Using the same notation as above, where in the case of recidivism at the legal limit, the outcome of interest, Y_{it}, is whether an individual, i, receives at least a second alcohol test with a BAC result of 0.08 or higher within some time period, t. T_{it} corresponds to the same treatment indicator as in the RDD; the indicator is equal to 1 if BAC $\geq$ 0.08, in the case of sanctions at the legal limit, and 0 if not. For the estimation of the effect of enhancements, the term T_{it} is equal to 1 if BAC $\geq$ 0.15, and 0 if not. A_{it} is a post-treatment indicator variable equal to 1 if the offense took place under the new law or 0 if not. In comparing the 2009 period to the 2005 period, $P_{it} = 1$ if the offense took place on or after March 24, 2009, when the date the law took effect, and 0 if it took place before that date. The model I estimate is the following:

$$Y_{it} = \alpha + \beta T_{it} + \gamma A_{it} + \delta T_{it} A_{it} + \varepsilon X + u_i. Y_{it} = tU\beta t_{it} + \gamma A_{it} + \delta AT_{it} T_{it} + \varepsilon X + u_{it}$$

The DID estimate is given by δ, which captures the average causal effect of the 2009 law on the outcome of interest. The matrix X includes fixed effects for year and county, in order to account for unobserved confounders, such as shocks unique to a particular year, the passage of legislation that might be correlated with the treatment in a specific a year, and local-level factors. u_i is a disturbance term, and clustered robust standard errors are used in all specifications. Standard errors are clustered at the county and year level to capture potential autocorrelation and heteroskedasticity. All specifications are for first-time offenders, and I run the specifications at various bandwidths in order to test the robustness of the results. Specifications are also run with

---

153. *See, e.g.*, Lee & Lemieux, *supra* note 90, at 285–86. As Caughey & Sekhon point out, a number of authors rely on a fourth-order (quartic) polynomial without justifying the functional form or presenting the results of alternative specifications. *See* Caughey & Sekhon, *supra* note 91, at 388, 397. We have no compelling *a priori* reason to choose a particular polynomial specification, which motivates the robustness check with varied specifications.

154. Veronica Grembi et al., *Do Fiscal Rules Matter?* 2 (Harvard Econ. Dep't Working Paper, Paper No. 397, 2015), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1852523.

second-, third-, fourth-, and fifth-degree polynomials to test the robustness of functional form. I run the same specification comparing the 2001–2005 period to the 2005–2009 period, where the local average treatment effect measures the effect of the 2005 law change that granted jurisdiction over IIDs to the Office of Driver Control.

APPENDIX II: ROBUSTNESS

A. Sensitivity to Bandwidth Choice and Model Specification

In order to examine the robustness of the RDD results, the bandwidth sensitivity plots below show the robustness of specifications to bandwidth and model choice. The models include difference in means models run for the key results in the paper.

**Figure 10: Bandwidth and Model Specification Sensitivity for
the Effectiveness of Sanctions on Recidivism at the
0.08 BAC Level During 2009-2013**

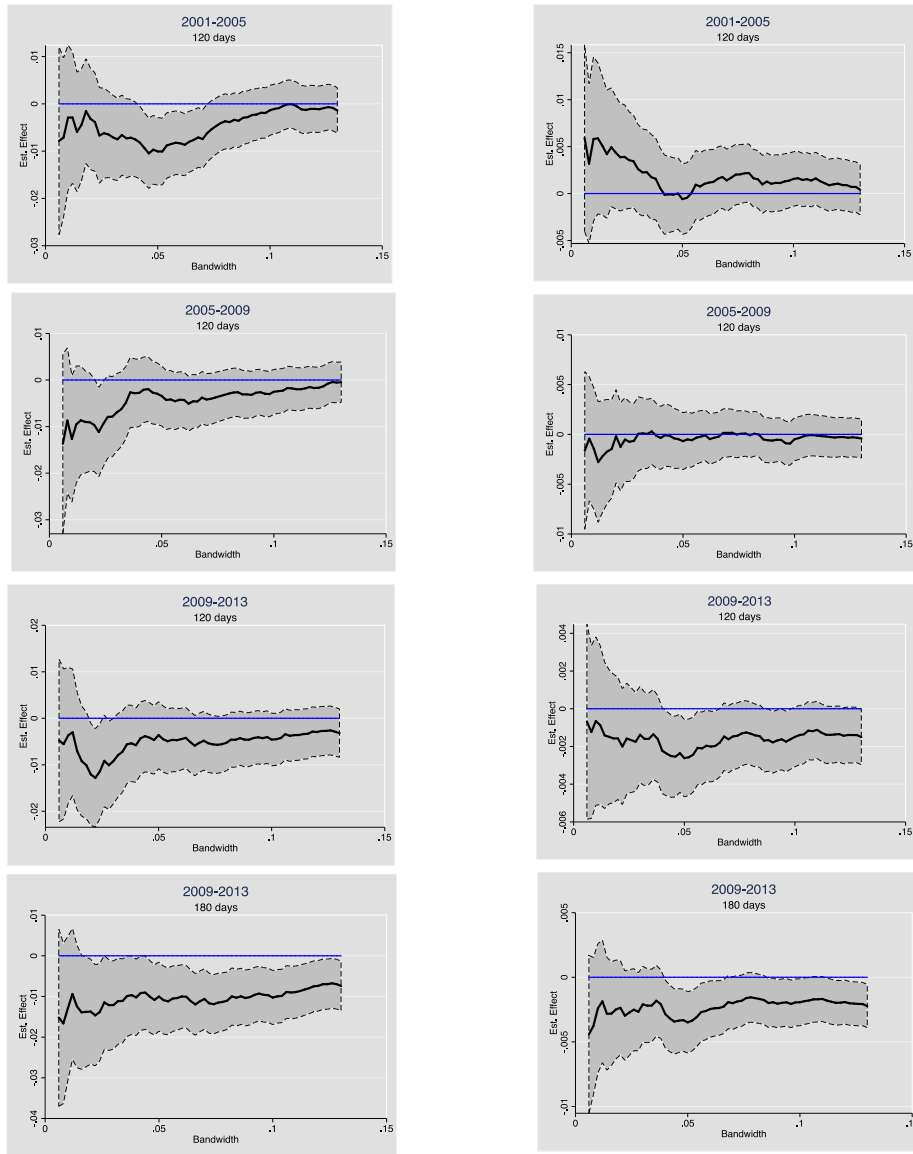Figure 10(a): Incapacitation and Recidivism   Figure 10(b): Incapacitation and Crashes

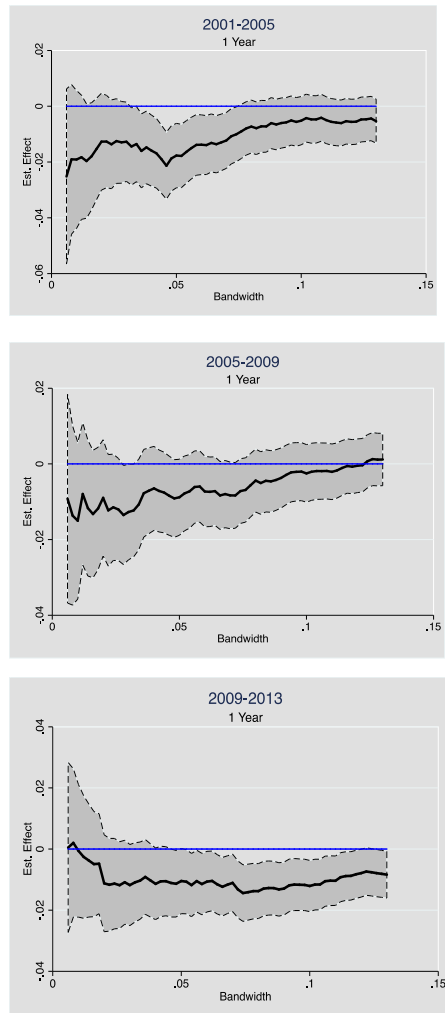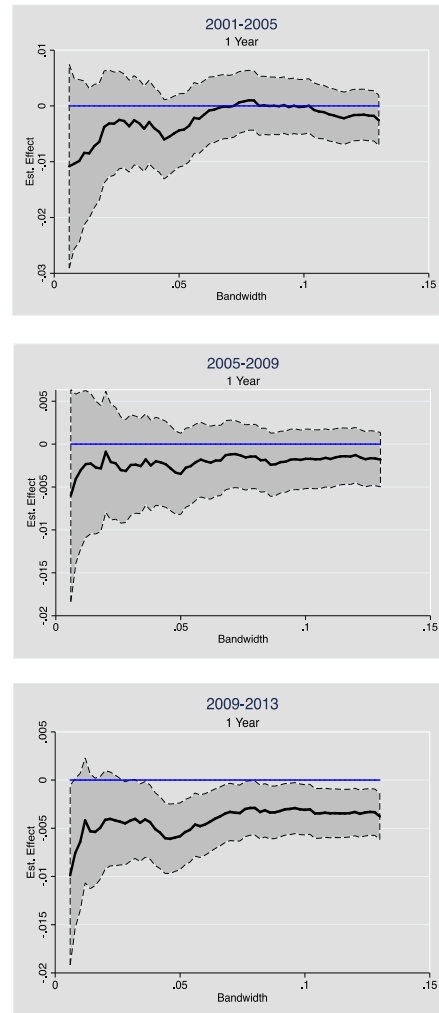Figure 10(a): Specific Deterrence and   Figure 10(b): Specific Deterrence and
            Recidivism   Vehicle Crashes

These plots show the sensitivity of estimates to bandwidth choice. The outcome is whether a first-time offender reoffended during a 180-day license suspension between over the important legal regimes that governed DWI in Arkansas from 2000 until the present.