The Iatrogenic Effects of Punishment

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Ryan Paul Larson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Chris Uggen
August 2022
Acknowledgements

In my rabid pursuit towards completing this dissertation, it was relatively easy to slip into the lull of isolated work and lose sight of the faculty, friends, and family members that have propelled me to finishing this piece of social science. I am grateful for the opportunity to express my immense gratitude to those who have helped me reach this point, but I worry I will fall short of truly expressing the debt I feel towards the people and institutions that have enabled me to pursue social science as a vocation.

I can say without much hesitation that my parents, Tim and Diane, have been, and continue to be, both my biggest supporters and my closest friends. I am really lucky to have been supported the way I have been throughout my time in school (and really, my entire existence), and it’s hard to see how I could have gotten this far without them. I’d also like to thank my brothers, Reid and Ross. Reid has always treated me like an expert in this stuff even if I speculated wildly at times, and Ross has fostered a musical relationship with me and has allowed me to honk my horn in the midst of his beautiful playing. I am further indebted to my grandparents, Marlene, Ron, Paul, and Mary, all of whom have served as models for how to approach life, vocation, and contentment, and I am grateful for their support I still benefit greatly from (and thanks to Grandma Marlene for watching all the Wild games with me).

My advisor, Chris Uggen, has been my most ardent supporter during my graduate career, and I’m not sure I can adequately express how much my scholarly thinking, practice, and focus is influenced by my time with him. I first learned of Chris from “coding up” his tweets as part of an undergraduate research project, and it’s weird how little things like that can truly alter the trajectory of your life. I am so thankful for Chris’ unwavering support, even in times where I felt I most certainly didn’t deserve it, as well as his good-natured friendship, collaborative mindset, and good humor that alone made my time in graduate school worthwhile. I feel like I have not only been trained in good science, but also how to be a generous collaborator, mentor, and person. I can’t recount here all the countless lessons I’ve learned from Chris over the years, but I will most certainly make sure to align the decimals.

I am also indebted to the members of my doctoral committee, Doug Hartmann, Josh Page, Ann Meier, and Aaron Sojourner. Doug brought me into the fold of the AMP project before I even set foot on campus at UMN, which gave me a bit of momentum to enter the program with some confidence and feel like I could make a contribution. Doug has had a consistent and thorough impact on broadening my sociological thinking, always encouraging me not to get lost in the details and remember how they can guide me to the big picture, on top of always being a really great person to be around. In my time working on DANCO cases with Josh, I was involved in my first foray into qualitative research, and this experience alongside his sociology of punishment class catalyzed my interests and opened my mind to punishment’s varied causes and effects, which I hope shows in this dissertation. Through our collaboration on the KIDS project, Ann helped me learn through some of the trials and tribulations of big data, and has shown me one can be both extremely kind and really, really smart all at once. I am also thankful to Aaron who mentored me on our paper on felony history an employment rates, pushed me
hard on my empirics and has exemplified how fruitful collaborations can be across disciplinary boundaries. I have learned so much from the time spent with each of them on research collaborations about the practice of social science during my time in graduate school, and I am grateful for their kindness and openness to working with me and giving me an opportunity to work alongside each of them.

My academic colleagues, friends, and collaborators are some of the most generous people I know. I am thankful to the elder members in the “Uggen shop”, who took me under their wing and always made me feel as though I was worthy of being part of their team. In particular, Rob Stewart, Veronica Horowitz, Sarah Shannon, and Sarah Lageson all have shown me respect, kindness, and friendship while working together throughout the years, and the time spent with them during collaborations and conferences has really been a blast. Thanks as well to Jeanie Santaularia, for the great collaborations and the help working through the hospital data. I’d also like to thank my 2014 cohort mates for the stimulating conversations we had in the office and good times outside of it that made graduate school pretty fun at times. Finally, I’d also like to thank the other UMN grad students who I have collaborated with, took classes with, worked on the TSP board with, or talked shop with: everybody’s passion towards the topics they were interested in drove me to think hard about my own.

I am also indebted to my undergraduate advisors and mentors, Andrew Lindner and Matthew Lindholm. I am thankful for their continued support and friendship, and my time spent with them at Concordia really allowed me to foster as a young social scientist. Lindholm first brought me on this path and gave me Seductions of Crime by Jack Katz the first time I visited his office, which was my baptism into sociological approaches to crime. Andrew really brought me into the world of quantitative social science at Concordia, and gave me opportunities and responsibilities that I would not have entrusted myself with. Andrew gave me my first copy of Stata and told me to “have fun!” Andrew has consistently been one of my closest collaborators, mentors, and friends, and he continues to be a model for how I hope to conduct myself as a professor. Their presence in my life at a liberal arts school inspired me to go on this journey, and hopefully I can do the same for students in the future.

I am grateful to the University of Minnesota Department of Sociology for having been given this opportunity to pursue a graduate degree. Amongst many reasons, one of the allure of going to the “U” was that I thought I would not just be relegated to a certain lane of academic inquiry, and that I wanted to become acquainted with the full scope the “big tent” of sociology has to offer. I’m happy to report that has indeed turned out to be the case. The department offered an environment to wrestle with and practice the craft of researching and teaching social science, and I have become a better sociologist, criminologist, teacher, and collaborator thanks to it. I’d also like to thank the department staff, especially Becky Drasin, for all the help navigating this process and life on campus.

My friends in the Twin Cities and elsewhere have provided an essential social network of friendship, fun, and support. Thanks to Kirk, Rob, Meg, Tom, Dani, Peary, Don, Fig, Peter, Jace, Emma, Molly, Jackie, Steve, Tyler, Jimmy, and many others for all the good times. I am truly at my happiest spending time with them, and for that I am grateful. Thanks are also due to my bandmates in both The Drug Budget and Sunshine &
The Nightwalkers. I appreciate their willingness to include me in their creative endeavors.

I’d also like to thank my new colleagues at Hamline University for being so welcoming and for taking a chance on me. I’m really excited to work and learn from them starting this Fall, and I will do my best not to let them down.

Finally, and certainly not least, thank you to my dog, Pig, for being such a good girl.
This dissertation is dedicated to my parents, Tim and Diane,

and my grandparents, Mary, Paul, Marlene, and Ron.

Thank you for everything.
Abstract

Sociological criminology has undergone a scholarly revolution in identifying the vast reach of punishment’s deleterious effects across multiple domains of American social, political, and economic life. However, this scholarship has largely neglected to empirically examine what ramifications these adverse effects of punishment have for crime. Across three empirical studies at multiple levels of analysis, this dissertation brings crime back in as a central outcome in the study of the effects of punishment, and examines aspects of the potential *iatrogenic*, or crime inducing, pathways of punishment. The first study, using court administrative data and the quasi-random assignment of judges in Minnesota, investigates the causal “packaging” effects of punishment on crime, finding the combination of hefty probation and monetary sanctions to be particularly criminogenic, alongside weak overall effects of punishment on recidivism. Second, this dissertation situates community-level punishment within sociological theories of neighborhood ecology and crime, and reveals *bifurcating* effects punishment on violence at the community-level, with incarceration and monetary sanctions loads tied to lower levels of neighborhood crime, but probation concentrations tied to higher crime rates. The second study also highlights a criminogenic path of punishment on violence by increasing neighborhood levels of concentrated disadvantage. The third empirical study leverages a difference-in-difference design to estimate the causal effect of ban-the-box legislation, which delays the disclosure of criminal records during the employment process, on both state-level employment and crime rates, finding little relationship between ban-the-box adoption and crime. In contrast, ban-the-box appears to bolster employment overall, but it may have adverse effects on Black employment. These empirical studies document the
iatrogenic links between punishment and crime, as well as examine the efficacy of state policy to sever these relationships.
# Table of Contents

Acknowledgements .......................................................................................................................... i

Abstract ........................................................................................................................................ v

Table of Contents .......................................................................................................................... vii

List of Tables .................................................................................................................................. xi

List of Figures ................................................................................................................................. xii

Introduction ...................................................................................................................................... 1

Background ...................................................................................................................................... 5

Organization of the 3-Chapter Dissertation .................................................................................. 9

Chapter 1: The Causal Packaging Effects of Punishment ................................................................. 12

Overview ........................................................................................................................................ 12

Introduction ...................................................................................................................................... 13

Literature Review ............................................................................................................................ 16

Theoretical Perspectives ................................................................................................................ 16

Incarceration and Recidivism ....................................................................................................... 17

Probation, LFOs, and Recidivism .................................................................................................. 20

Packaging of Punishment ............................................................................................................. 21

Current Chapter ............................................................................................................................. 22

Data and Measures ....................................................................................................................... 24
UKCPR National Welfare Data................................................................. 125
Federal Bureau of Investigation Uniform Crime Reports......................... 125
Felony History Shares............................................................................. 126
Bureau of Justice Statistics – National Prisoner Statistics Survey ............... 127
Analytical Strategy................................................................................... 128
Descriptive Analyses of BTB, Employment, and Crime......................... 128
Difference-in-Difference with Staggered Adoption .................................. 130
Results...................................................................................................... 139
Descriptive and Bivariate Results........................................................... 139
Staggered Adoption Difference-In-Differences....................................... 145
Discussion and Conclusion..................................................................... 154
Conclusion: Sociological Insights, Policy Implications, and Future Directions 161
Bibliography ............................................................................................ 167
List of Tables

Table 1 Descriptive Statistics for Variables in Full Risk Set.................................................. 30
Table 2 Packages of Punishment .................................................................................................. 42
Table 3 IV 2SLS Models of Recidivism...................................................................................... 47
Table 4 IV 2SLS Models of Recidivism - Incar. X Prob................................................................. 50
Table 5 IV 2SLS Models of Recidivism - Prob. X LFO ................................................................. 52
Table 6 IV 2SLS Models of Recidivism - Incar. X LFO................................................................. 53
Table 7 IV 2SLS Models of Recidivism - Pub. Def. Interactions .................................................. 55
Table 8 Descriptive Statistics for Panel....................................................................................... 80
Table 9 CFA Measurement Model of Concentrated Disadvantage.............................................. 88
Table 10 FE Panel Models of Concentrated Disadvantage, 2011-2014 ....................................... 92
Table 11 FE Panel Models of the Violence Rate, 2011-2014....................................................... 94
Table 12 BTB State-Level Legislation .......................................................................................... 123
Table 13 Descriptive Statistics for All Variables ......................................................................... 128
Table 14 Focal DV Means by Pre-Post BTB .............................................................................. 145
Table 15 Staggered Adoption DID Conditional Aggregated Group-Time ATTs of First BTB Law....................................................................................................................... 153
List of Figures

Figure 1 IV Model of Recidivism using Random Judge Assignment .......................... 32
Figure 2 First-Stage Judge Effects .............................................................................. 37
Figure 3 Recidivism Events by Highest Reconvicted Charge .................................... 43
Figure 4 Recidivism Events by Punishment Package .................................................. 45
Figure 5 Probation X LFO Interaction Plot ................................................................ 52
Figure 6 Incarceration X Public Defender Interaction Plot ...................................... 56
Figure 7 Theoretical Model of Punishment, Disadvantage, and Violence ................. 74
Figure 8 Violence Rates and Punishment Loads by ZCTA ........................................ 87
Figure 9 Mediation Analysis of Incarceration ............................................................ 98
Figure 10 Mediation Analysis of Probation ................................................................. 98
Figure 11 Mediation Analysis of LFOs ......................................................................... 98
Figure 12 State-Level BTB Legislation, 2020 .......................................................... 140
Figure 13 Kaplan-Meier Survival Plot of BTB Legislation .................................... 141
Figure 14 State Prime-Age Nonemployment Shares, 2020 ....................................... 142
Figure 15 State Index Crime Rates, 2020 ................................................................. 142
Figure 16 Nonemployment Shares by BTB Legislation ............................................. 144
Figure 17 Total Index Crime Rates by BTB Legislation ............................................... 144
Figure 18 Staggered Adoption DID Group-Time ATTs of First BTB Law .................... 147
Figure 19 Staggered Adoption DID Group-Time ATTs of Public BTB+ .................... 149
Figure 20 Staggered Adoption DID Group-Time ATTs of Public and Private BTB .... 151
Figure 21 Staggered Adoption DID Group-Time ATTs of First BTB Law .................. 153
Introduction

This dissertation arises from questions surrounding the relationships between punishment and crime in the context of a rise in the use of multiple different forms of punishment in the United States, alongside a general decrease in crime and victimization. The massive scaling up of the criminal legal apparatus from 1975-2020, combined with the substantial decrease in the crime rate, may lead one to believe that the particularly palpable use of punishment has causally led to this marked increase in public safety in American society. However, substantial theoretical and empirical scholarship within the sociology of punishment has complicated this simplistic explanation that crime is a simple function of punishment, nor that punishment is merely responsive to fluctuations in reported crime rates (e.g., Durkheim 1893; Rusche and Kircheimer 1939; Foucault 1975; Garland 1993; Beckett 1999; Wacquant 2009). In the words of Rusche and Kirchheimer (1939),

“The bond, transparent or not, that is supposed to exist between crime and punishment… must be broken. Punishment is neither a simple consequence of crime, nor the reverse side of crime, nor a mere means which is determined by the end to be achieved. Punishment must be understood as a social phenomenon freed from both its juristic concept and its social end (5).”

Thus, the link between punishment and crime is complex, and multiple institutional, economic, political, cultural, and social forces are complicit in the patterning of both punishment and crime.

In this vein, contemporary sociological punishment scholarship largely eschews crime as a focal outcome of interest in relation to the effects of punishment at both the
individual and aggregate levels. Early studies in modern criminology focused primarily on the question of crime, investigating how punishment might best be used as a crime control device (Sampson 2011). However, criminological work within sociology has shifted to focus on the effects of punishment on other facets of social, economic, and political life. This shift has proven fruitful in identifying the consequences of punishment across a broad swath of social life from employment (Pager 2003; Pager, Bonikowski, and Western 2009; Uggen et al. 2014; Larson et al. 2022), income and earnings (Western 2002), health (Schnittker, Massoglia, and Uggen forthcoming; Schnittker et al. 2011; Schnittker et al. 2015; Sugie and Turney 2017), housing (Harding et al. 2013), family life (Comfort 2008), child well-being (Wakefield and Wildeman 2013), voting and political participation (Manza and Uggen 2008; Sugie 2015; Uggen et al. 2020), education (Stewart and Uggen 2020), and institutional engagement (Brayne 2014) to name but a few. In effect, the powerful research on collateral consequences (Kirk and Wakefield 2018; Uggen and Stewart 2014) has highlighted the vast social harms that punishment has wrought during the era of mass incarceration. Relatedly, these harms are in addition to, or perhaps exacerbated by, the racialized social ills of aggressive modern policing (Kramer and Remster 2022; Rios 2011; Edwards et al. 2019; Legewie and Fagan 2019).

Thus, the majority of the recent scholarship within sociological criminology interrogates the insidious supplementary effects of punishment without much talk of what implications punishment’s noxious properties has for crime. Sampson (2011) has defined these two major eras of crime and punishment scholarship respectfully as crime control and crime production, with the first era being marked by investigation of the social influences on crime etiology, deterrence theories, and incapacitation effects of
punishment, and the second characterized by an intense focus on the multitude of collateral consequences and negative consequences of punishment. While a vast amount of research indicates the causal relationship between punishment and crime is not a simple one-to-one corollary, punishment still has both direct and indirect implications for crime. As Sharkey (2018b) astutely notes, while empirical evidence suggests that the penal expansion and hyper-aggressive policing characteristic of modern American punishment has played a role, albeit a modest and grossly inefficient one, in the great American crime decline (e.g., Johnson and Raphael 2012; Worrall and Kovandzic 2010), these gains have been accompanied by a vast array of social ills that have aided in the maintenance of social stratification and have devastated disadvantaged communities.

This logic is closely related to what Sampson (2011:824) calls the “social ledger of incarceration” – which entails empirically detailing “the full ramifications of incarceration’s costs and benefits.” Sampson (2011) discusses the oft complex and sometimes bifurcating effects punishment can have on both crime and various aspects of social life, and he calls for studies that “attempt to sort out the kinds of complexities inherent to incarceration (824).” This incarceration ledger is exemplified in Light and Marshall (2018), who compare the returns to incarceration in terms of reductions to homicide to the increased infant mortality rates resultant from high levels of incarceration. They find that, when compared on a common metric, lives, the benefits of incarceration are wildly overstated (and may even be null) when one considers in tandem the lives lost in terms of infant mortality to this punitive practice. Sampson (2011) states that research in this vein is indicative of a third era in the sociological study of crime and punishment, where the complex and varied effects of punishment on crime are under
consideration. Further, Sampson (2011) stresses that the accounting of punishment’s complex causal constellations is essential for social change and public policy, and research that neglects these complexities are “problematically related to policy recommendations (826).”

Leaving crime out of the equation in the investigation of punishment effects is inattentive to the ways in which punishment can *exacerbate* crime and criminogenic contexts, both directly and indirectly through its impact on existing stratification and collateral consequences. In other words, contemporary institutions of social control, such as punishment regimes, may, via their vast impact on the social lives of those under its purview, strengthen or promote the very social problem its methods and tactic purport to heal. Questions remain as to how punishment’s varied forms, intensities, and contexts are implicated in the (re)production of class and racial inequalities (Wakefield and Uggen 2010), the structuring of American neighborhoods and contexts (Sampson 2012), and corrosion of general social health (e.g., employment); all of which have been shown in previous research to be associated with *elevated or perpetuated levels of crime*. In sum, punishment may not only produce crime in the Foucauldian (1975) discursive sense of the social construction of knowledge, categorizations, and disciplinary practices surrounding criminal behaviors, but may also directly recreate the social conditions necessary for the genesis of crime.

Thus, this three-paper dissertation situates itself firmly in this third era of sociological crime and punishment research, empirically investigating the complex relationship between crime and punishment across multiple levels of analysis. Iatrogenesis is the phenomenon of a medical treatment, applied to an illness, that has effects that are
antithetical to treating the malady at hand, and has been used by Joan McCord (McCord 1978; Trembley et al. 2019) to describe counterproductive delinquency-prevention programs. I borrow this term and apply it to the sociological study of punishment to represent causal pathways of punishment that can lead back to perhaps increased, or maintained, levels of crime. In concert with scholarship that investigates the varied direct and indirect effects of punishment (e.g., Light and Marshall 2018), this dissertation contains three empirical chapters that speak to the direct and indirect iatrogenic effects punishment can have on crime, as well as evaluate a potential policy mechanism that has the potential to sever an iatrogenic causal link between punishment and crime. Further, I extend Sampson’s (2011) notion of the “incarceration ledger” to the “punishment ledger”, by extending this framework to non-custodial forms of punishment such as community supervision and legal financial obligations (LFOs). This dissertation will bring crime back in as a central outcome measure in the study of punishment effects, as well as expand upon the existing body of sociological and criminological research by using novel measures of punishment, unique causal identification strategies, and testing extensions to criminological theory to further elaborate pieces of the punishment puzzle.

Background

Since the mid-1970s, the United States has experienced a substantial rise in incarceration and other forms of punishment. The imprisonment rate – the number of individuals in prisons per 100,000 adult population – was 161 in 1972, and peaked in 2007 at 670, and has decreased modestly to 358 at year end 2019 (Carson 2020). This marks the use of incarceration in the United States as exceptional in both a historical and international comparative sense (Garland 2001). Concurrent with the rise in mass
incarceration has been the rise in mass probation (Phelps 2013), but the latter has taken a slightly different spatial and demographic trajectory (Phelps 2017). About 1 in 58 US adults is under community supervision – probation or parole (Kaebble and Alper 2020) – which represents a far larger correctional population pool than those incarcerated, with the probation to incarceration ratio per 10 index crimes at 2.7 in 2010 (Phelps 2013). Recent inquiry has also established the prevalence of Legal Financial Obligations (LFOs) as a significant mode of punishment (Harris 2016), that shows substantial variation by place and could play a significant role in the reproduction of inequality (Martin et al. 2018; O’Neill et al. 2021). Millions of Americans also experience frequent police contact (Stuart 2016), as well as the procedural hassles that are inherent amongst misdemeanor-level cases (Kohler-Hausmann 2013). These increases in punishment have conspired to mark a substantial slice of the American population with a criminal record. Shannon et al. (2017) estimate that 5.1 million former prisoners currently reside in the United States, which translates to 6% of all adult men and 15% of African American adults that are subject to the mark of a felony criminal label.

Extensive research has shown that experience of these ever-prevalent forms of punishment is patterned by race and class, as more than half of African American men with less than a high school degree will go to prison at some point in their lifetimes (Pettit & Western 2004). Similar racially disproportionate patterns are seen in police contact (Gelman et al. 2007), arrest (Brame et al. 2012; Brame et al. 2014), sentencing (Johnson 2003; Harris et al. 2011; Kutateladze et al. 2014), and criminal records (Shannon et al. 2017), amongst other domains. From the racial patterning in mortality from police brutality (e.g., Edwards et al. 2019), to the predatory resource extraction of
poor communities through monetary sanctions and other mechanisms (Harris 2016; Page and Soss 2021), this work highlights how the negative consequences of punishment are also patterned along race and class lines.

Researchers have attributed such racial disparities to mechanisms in both the criminal legal system, and broader society as a whole, such as overt racial discrimination in decision making and policing, political fear-mongering along the lines of race, as well as more insidious institutions of structural racism in society, such as residential segregation and racial stratification (Peterson and Krivo 2010; Alexander 2012; Forman 2017; Wacquant 2000; Wacquant 2001). This conceptualization of the modern punitive regime indicates that “collateral” consequences are not really all that “collateral” - mere latent byproducts of punishment’s application-, but rather are part and parcel of projects of social control, surveillance, and discipline directed towards racial minorities and impoverished populations (Hinton and Cook 2021; Wacquant 2009; Soss et. al. 2011; Rusche and Kirchheimer 1963). In sum, scholarship has indicted punishment as a part of the toolkit of racialized social control, disciplinary governmentality, predatory extraction, and oppression.

This racialized rise in punishment of multiple varieties has coincided with a decrease in the national crime rate (Zimring 2006). Both official arrest data (Uniform Crime Report 2019) and the National Crime Victimization Survey (Morgan and Thompson 2020) show a marked drop in crime since 1991, after a steady rise from the 1960s. Despite this national decrease in crime, scholarship has expanded upon the adverse consequences of exposure to crime and violence in both direct and indirect forms that go far beyond immediate victimization. A breadth of surveys highlights that although the
prevalence of violence has declined, violence remains a relatively common experience in the United States, with particularly higher exposures for youth and racial minorities (Sharkey 2018a) and a spatial concentration in social and economically disadvantaged neighborhoods (Braga et. al. 2010).

Evidence shows that this concentration of violence and violent victimization are partly a consequence of patterns in racial segregation (Krivo et al. 2009; Peterson and Krivo 2010), as well as the resulting concentrated disadvantage and other enduring inequalities of neighborhood context (e.g., Peterson and Krivo 2005; Sampson 2012). Further, these structural effects on violence may be stable across racial groups, and it is the unequal exposure of individuals to these structural effects that pattern crime and violence (Krivo and Peterson 2000). This enduring and racially stratified exposure to crime and violence is not only criminogenic, but is also tied to children’s cognitive performance (Sharkey 2010), school testing (Sharkey et. al. 2014), and subsequent gunshot victimization (Papachristos et. al. 2015). Further, some of these negative consequences of exposure to crime have been shown to be exacerbated for young, minority individuals (e.g., Sharkey et. al. 2014; Papachristos et. al. 2015).

Concentrations of violence, along with other racial inequalities of place, can feedback into the political economy of opportunity structures and investment (or lack thereof) at both individual and community levels, which can help reproduce levels of crime and disadvantage (Sharkey 2013; Squires and Kubrin 2005). In sum, while crime has declined overall over the past decades (despite modest recent upticks in violent crime in certain metropolitan areas (Rosenfeld and Lopez 2021)), the effects of criminal victimization are shown to be important for the social health and functioning of communities.
Organization of the 3-Chapter Dissertation

The first empirical chapter of this dissertation investigates the direct effects of punishment on crime at the individual-level of analysis. Extensive previous research has examined the link between punishment and recidivism in the context of an isolated mode of punishment, primarily looking at the impact of incarceration (e.g., Harding et al. 2017). However, not only are the different forms of punishment (incarceration, probation, monetary sanctions) potentially related to one another, but they are also experienced together. Thus, the first empirical chapter sets out to analyze the “packaging,” or conditional, effects of punishment and examine punishment more closely as it is experienced by defendants. This will not only assess the relative impacts of individual-level punishment in tandem, thereby avoiding potential model misspecification in previous work, but also extend this literature by investigating how packaging punishment types together may interact to pattern crime. Using a robust court administrative dataset of criminal punishment in Minnesota, alongside an empirically rigorous causal identification strategy leveraging the quasi-random assignment of judges, the first dissertation chapter will speak to the iatrogenic effects of punishment at the individual-level.

The second empirical chapter situates punishment within sociological neighborhood ecology theories of crime, which contend that neighborhood structure translates to neighborhood-level crime via the structuring of informal social controls, collective efficacies, and social networks in communities (e.g., Sampson 2012). These theories have largely neglected to consider what role neighborhood-level concentrations of punishment play in these ecological models of crime. In concert with scholarship that
identifies punishment as a key driver in the (re)production in American stratification (e.g., Western 2006), I ask to what extent punishment may pattern aggregate neighborhood-level crime through the (re)structuring of neighborhood levels of disadvantage – a potential iatrogenic pathway from punishment to crime at the meso-level. Using a uniquely constructed panel dataset comprising aggregated court administrative data and hospital discharge records in Minneapolis, MN, this empirical chapter analyzes a) the direct effect of community concentrations of different forms of punishment on crime, and b) the indirect effect of punishment concentrations on neighborhood-level violence via the mediating terrain of community disadvantage.

The United States has been shown to be particularly proficient in the production of felony-level criminal records, with over 19 million Americans now subject to a felony-level criminal record (Shannon et al. 2017). Such records could serve as a potential iatrogenic causal pathway by which punishment may translate to increased levels of crime. For example, the collateral consequences literature has been adept at investigating the impact of criminal records on employment across multiple levels of analysis (e.g., Pager 2003; Larson et al. 2021), as well as documenting the salience of work for desistance from crime (e.g., Uggen 2000; Sampson and Laub 1993). In sum, the state-level production of criminal records, coupled with the penchant for employers to consider this information, represent an iatrogenic causal chain by which punishment may pattern aggregate crime rates. Thus, the third empirical chapter in this dissertation examines a mechanism within the realm of state-level public policy with the potential to sever this iatrogenic link: ban-the-box legislation. This chapter uses a generalized difference-in-difference design to estimate the causal relationship between ban-the-box legislation and
subsequent state-level employment, and also examine the heterogeneity in the impact amongst different law types and demographic groups, as previous economic research has suggested that these laws may hurt the employment prospects of Black Americans (e.g., Doleac and Hansen 2020). This chapter also directly tests the ability of ban-the-box legislation to help criminal record holders by examining the impact in relation to states’ criminal record production. Finally, the chapter extends the scholarship on ban-the-box legislation by investigating to what extent ban-the-box legislation severs the iatrogenic relationship by impacting state-level crime rates.

Overall, this dissertation makes contributions to the sociological study of crime and punishment by exploring the potential iatrogenic impacts of punishment on crime, keeping a keen eye on how punishment structures social life more broadly, and often in a racialized manner. The subsequent three chapters will discuss each of the empirical studies just described, including the relevant bodies of existing scholarship, data construction and research design, analytical strategy, statistical assumptions, results, and the implications the findings hold for both sociological research and public policy. After these empirical chapters, I close with a concluding chapter that discusses the overarching implications for this research, as well as future directions in the sociological study of the “social ledger of punishment.”
Chapter 1: The Causal Packaging Effects of Punishment

Overview

The United States has experienced a marked rise across multiple forms of punishment, such as incarceration, probation, and monetary sanctions during the era of mass incarceration. However, studies examining the impacts of punishment on recidivism have a) primarily focused solely on the effects of incarceration and b) treated each axis of punishment as a separate, isolated factor, in contrast to how they are experienced as a “package” of punishment. This study leverages a natural experiment using the quasi-random assignment of judges within Minnesota Judicial Districts and estimates the causal impacts of incarceration, probation, and monetary sanctions on recidivism, while also examining how each punishment mode may condition the impact of the other. Findings indicate that the relationship between punishment and subsequent crime is weak overall, but the packaging of probation and LFOs together may be criminogenic. Further, negative effects for incarceration are only present amongst more advantaged defendants. These findings suggest that the impacts of punishment should be studied as they are experienced – as a mutually-constitutive package of confinement, surveillance, and financial costs.
Introduction

While the United States has indeed experienced a steep rise in incarceration from the 1970s to the late 2000s of the mass incarceration era (Garland 2001; Carson 2020), this increase coincided with rises in the use of community supervision (Phelps 2013; Phelps 2017) - such as probation – and predatory financial extractive practices such as monetary sanctions such as fines, fees, and restitution (Page and Soss 2021; Harris 2016). Thus, the mass incarceration era is not only characterized by an increase in incarceration rates, but also the rising use of both community supervision and a variety of financial fines and fees; an era that might be more appropriately called the era of “mass punishment.”

This wave of criminal punishment has also been strongly racialized in the United States. The steep rise in jail and prison populations is really the story of racialized mass incarceration, with Black men experiencing especially high rates of imprisonment and felony conviction (Pettit and Western 2004; Shannon et al. 2017), as well as being more likely to receive confinement sentences as opposed to probation sentences (Freiburger and Hiliniski 2013; Harrington and Spohn 2007). Further, probation sentences amongst racial groups have been shown to be unequal for similarly situated cases (Lowder et al. 2019), and a similar racialized pattern holds true for the imposition of monetary sanctions as well (Harris et al. 2011). This racialization in mass punishment has also led to the unequal proliferation of felony-level criminal records amongst the U.S. population, as Shannon et al. (2017) estimates that 8% of Americans as of 2010 had a felony conviction, and 33% of African American men held a felony record. Researchers have attributed such racial disparities to mechanisms in both the criminal legal system and broader society as a
whole, such as overt racial discrimination in decision making and policing, political fear-mongering along the lines of race, as well as more insidious institutions of structural racism in society, such as residential segregation and racial stratification (Peterson and Krivo 2010; Alexander 2012; Wacquant 2000). This scholarship calls into the question the stated aims of the punishment apparatus as a mechanism for public safety, and instead frames punishment as a part of the toolkit of racialized social control and oppression (Hinton and Cook 2021).

An essential and enduring question within criminology concerns whether punishment has an impact on the likelihood of crime commission in the future at the individual level. These questions are particularly salient in a time when the amount of punishment experienced per defendant has risen, and when the efficacy of the criminal legal apparatus to enhance public safety is under intense academic scrutiny (Cullen et al. 2011; Alexander 2012). Various criminological theories, such as deterrence and labeling theories, make different predictions as to how punishment will impact individuals’ likelihood of reoffence based upon sentenced amounts of punishment. Further, the extensive empirical work on the collateral consequences of punishment (e.g., Uggen and Stewart 2015) complicate the relationship between punishment and crime, and represent potential crime-promoting influences of punishment that may offset any gains made via punishment’s more classic assumed causal pathways.

A wide breadth of studies examines the impacts of incarceration on recidivism and a variety of other outcomes (e.g., Harding et al. 2017; Harding et al. 2018), generally finding null or weak criminogenic effects at the individual level that can vary based on defendant levels of disadvantage and other social factors (e.g., Mears et al. 2015). A
fewer number of studies have examined the effects of probation and monetary sanction loads on recidivism. Despite these empirical studies, few studies to date have examined the effects of each form of punishment in tandem, as the imposition of one form of punishment could confound the relationship between another and recidivism. Further, studies have neglected to properly assess not only the independent effects of each form of punishment net of the other types, but also examine how these varying forms and intensities of punishment combine to pattern recidivism outcomes. Thus, questions remain as to how the “packaging” of punishment – the constitutive amalgamations of variable punishment kinds and concentrations – pattern recidivism outcomes amongst those sentenced. I see this as a fundamental sociological insight largely ignored within the sociological and criminological study of punishment and recidivism, in that punishment should, and can, be analyzed more closely as it is experienced – as a mutually-constitutive package of confinement, surveillance, and financial penalties.

Therefore, this empirical dissertation chapter asks three related research questions: 1) To what extent do sentenced amounts of incarceration, probation, and monetary sanctions affect recidivism? 2) Does the “packaging” of these punishments together alter their causal impact? And 3) To what extent is there treatment heterogeneity in punishment amongst levels of defendant disadvantage? To answer these empirical research questions, I follow in a line of rigorous empirical scholarship that uses quasi-experimental designs to deal with the endogenous nature of crime and punishment (e.g., Harding et al. 2017) to estimate the causal packaging effects of punishment on recidivism. Using court administrative data from the state of Minnesota, I leverage the quasi-experimental assignment of judges to felony-level cases to estimate the causal
effects of varying punishment intensity on recidivism. Generally, I find weak causal effects of each form of punishment on recidivism, with varying effects of incarceration conditional on defendant disadvantage. However, I identify the confluence of probation and monetary sanctions as having a particularly salient criminogenic interaction. These findings extend the sociological study of punishment and recidivism by leveraging a rigorous empirical design to assess the packaging effects of varying forms and intensities of punishment, and highlighting the conditional, and ultimately underwhelming, nature of punishment’s effects.

**Literature Review**

*Theoretical Perspectives*

Deterrence theories propose that increases in punishment intensity, certainty, and celerity will decrease the likelihood of future crime, with mechanisms such as incapacitation (Piquero and Blumstein 2007), rehabilitation (Lipsey and Cullen 2007) and specific deterrence (Nagin et al. 2009; Nagin 2013) decreasing the likelihood of future offending. Incapacitation effects occur while an individual remains incarcerated, and specific deterrence and rehabilitation effects occur post-release (Harding et al. 2017). In contrast, labeling theory (Lemert 1951) posits that punishment leads to increases in subsequent criminal conduct through three related channels: alterations of self-conceptions (Matsueda 1992), devolution of social relationships and opportunities (Becker 1963; Sampson and Laub 1993), and increased social control and surveillance. For example, Chiricos et al. (2007) find that those with adjudication withheld (i.e., when a conviction is withheld for a period of time) are significantly less likely to recidivate as compared to comparable defendants whose adjudication was not withheld.
Relatedly, the recent scholarship on collateral consequences suggests that punishment disrupts various aspects of social life such as employment (Pager 2003; Uggen et al. 2014; Harding et al. 2018), education (Stewart and Uggen 2020; Bernburg and Krohn 2003), and voting (Manza and Uggen 2008; Uggen et al. 2020) that may have a positive relationship to desistance from crime. This scholarship suggests that criminal punishment can alter aspects of the life course (Massoglia and Uggen 2010) that are integral for desistance from crime. This marks a crime-inducing causal pathway of punishment to crime, as criminal justice contact can institute barriers to social integration and desistance.

Thus, punishment may, through different mechanisms, have bifurcating effects on future crime, and looking at the total effect of punishment is a combination of each of these effects. However, their relative effect sizes determine whether this cumulative effect will result in a net increase or decrease in subsequent crime. In sum, criminological theory is suggestive of two bifurcating pathways that may work to cancel out the other in terms of the overall effect, and empirical studies generally estimate the net effect of a form of punishment on subsequent recidivism.

Incarceration and Recidivism

Empirical research on the impacts of punishment on recidivism has primarily examined the effects of incarceration. An extensive empirical literature based on analyses of individual-level observational data has generated mixed conclusions, with some studies finding that those who are given longer incarceration sentences are less likely to recidivate (Smith and Garton 1989), more likely to recidivate (Spohn and Holleran 2002), or that punishment has no impact on subsequent reoffence (Gottfredson 1999) or varies
by whether recidivism is measured by a new sentence or a technical violation (Rydberg and Clark 2016). Further observational research suggests that the effects of incarceration length on crime may also bleed into outcomes for children of those incarcerated (Mears and Siennick 2016). Recidivism following incarceration in a supermax facility has also been studied, finding an increase in violent recidivism post-release as compared to incarceration in non-supermax facilities (Mears and Bales 2009).

The overwhelming majority of individual-level observational studies find either no association between the experience of imprisonment and subsequent offending or a slight criminogenic effect (Nagin et al. 2009). However, these studies, which primarily base their causal identification on covariate adjustment via regression approaches with observational data, are susceptible to bias insofar as the unobserved attributes of a defendant or case that are associated with sentencing also might affect the defendant’s probability of recidivism. In other words, individuals sentenced to longer incarceration sentences may be systematically different from those who are not in ways that are unobserved by researchers. Therefore, there is the threat of omitted-variables-bias to the internal validity of estimates derived with these analytical approaches from observational data.

Studies using quasi-experimental designs to account for the endogeneity of punishment and recidivism generally corroborate the findings of observational studies with covariate adjustment (Killias et al. 2000; Green and Winnik 2010; Loeffler 2013; Nagin and Snodgrass 2013; Mueller-Smith 2015; Harding et al. 2017; Williams and Weatherburn 2022; Loeffler and Nagin 2022; see Roach and Schanzenbach 2015 for a negative effect). In one of the earlier studies leveraging a random judge IV design, Green
and Winnik (2010) find no effect of incarceration on rearrest using a 4-year follow-up window. Similarly, Loeffler (2013) finds no significant incarceration effects on either recidivism or employment in a study of defendants in Cook County, Illinois. Nagin and Snodgrass (2013) also find no effect of imprisonment on subsequent crime.

Other quasi-experimental studies have identified a criminogenic effect of incarceration on recidivism. Mueller-Smith (2015) finds an increase in recidivism in a Texas study, and Harding et al. (2017) find that incarceration, in comparison to probation, leads to an 18 percent increase in the likelihood of subsequent imprisonment amongst non-White defendants and a 19 percent increase among White defendants. Studies have also found similar effects among juveniles sentenced to incarceration (Aizer & Doyle 2015). Regression continuity designs, which causally identify the effect of incarceration by comparing individuals at some scored cut-off point (e.g., sentencing grids, other screening tools), generally corroborate the findings from the quasi-experimental IV designs (Loeffler and Nagin 2022; Mitchell et al. 2017; Franco et al. 2020; see Rose & Shem-Tov 2020 for negative findings).

The breadth of these studies has led scholars to conclude that incarceration does not have a specific deterrence effect, and the research evidence is more in line with a criminogenic effect of incarceration (Loeffler and Nagin 2022; Cullen et al. 2011). However, it should be noted that these designs generally use a non-incarcerated group (i.e., those sentenced to probation) as a comparison group, thereby estimating the effect of the presence of incarceration, at any amount, to defendants sentenced with probation. Studies that specifically compare the intensity of incarceration, although less numerous in
the literature, find weak negative effects as a result of longer incarceration sentences (Harding et al. 2017, Roach and Schanzenbach 2015).

**Probation, LFOs, and Recidivism**

Comparatively less research has examined the recidivism impacts of probation and monetary sanctions as compared to that of incarceration. A corollary of the studies comparing incarceration to probation generally finds that those sentenced to probation in lieu of a confinement sentence recidivate at lower levels (e.g., Harding et al. 2017). Additionally, research on the effects of probation indicate that different types of probation (e.g., in-home vs. out-of-home) can impact subsequent recidivism likelihoods (Ryan et al. 2014; Piquero 2003). Using quasi-experimental methods, probation lengths have been shown to have no effect on subsequent likelihoods of recidivism (Green and Winnik 2010).

In addition to the increase in the use of probation (Phelps 2013), legal financial obligations (LFOs) have also become more prevalent forms of punishment in the U.S. (Harris et al. 2010; Harris 2016; Martin et al. 2018). The limited observational evidence (for a review see Link 2022) shows mixed findings for financial penalties in terms of its impact on recidivism. Monetary sanctions may reduce the likelihood of reoffence, particularly in terms of the impact of restitution (Gordon and Glaser 1991; Ruback et al. 2004; Schneider 1986; Butts and Schnyder 1992). Conversely, a study of fines levied against drunk-drivers in Australia found no effect on subsequent recidivism (Weatherburn and Moffatt 2011). Studies also find mixed effects of monetary sanctions on recidivism depending on the type of LFO and how recidivism is measured. For example, Ruhland et al. (2020) find criminal fees associated with higher likelihood of
probation revocation but find no relationship for fines and a negative relationship for restitution. Consistent with a somewhat mixed finding, Iratzoqui and Metcalfe (2017) find that probationers receiving financial penalties are no more likely than those not imposed financial sanctions to garner a probation violation, but the violations amongst those with LFOs tended to be more severe.

Studies also suggest that LFOs may have a criminogenic effect on defendant recidivism. Using longitudinal national register data from Finland, Aaltonen et al. (2016) find that crime is higher among individuals during periods of debt. A study of a cohort of juveniles in Pennsylvania found LFOs increase the odds of 2-year recidivism by 33%, and this effect was greater amongst those levied higher financial LFO amounts (Piquero and Jennings 2017). Further, evidence from defendant interviews indicate that the imposition of LFOs can increase contact and surveillance, and defendants may resort to illegal means of “quick cash” to help pay off criminal legal debt, thereby signaling a criminogenic influence of LFOs (Horowitz et al. 2022; Harris et al. 2010; Ortiz and Jackey 2019). In sum, the research on LFOs and recidivism is decidedly mixed, and no study to date has adequately accounted for selection bias concerns in estimating the impact of LFOs on recidivism.

Packaging of Punishment

Recent research suggests not only that these modes of punishment are meted out in tandem (Harris 2016), but also that they are correlated with each other (Waldfogel 1995; Martin 2012). Specifically, Martin (2012) finds that the probability of a monetary penalty is negatively associated with the likelihood of an incarceration sentence, but the severity of the prison sentence is positively associated with monetary punishment. In
other words, financial punishments may serve as both a substitute and compliment to carceral forms of punishment, depending on the severity of the prison sentence. Further, the packaging of punishment a defendant receives, for similarly situated cases, is patterned by the defendant’s race (Larson et. al. working paper). Statistically, because amounts of different parts of punishment are associated, analyses of the impact of a singular aspect of punishments may be mis-specified due to the dependent nature of the different forms of punishment. Further, this suggests that quasi-experimental estimates in the literature may violate the exclusion restriction, as judge instruments have an alternative causal pathway to recidivism through alternative modes of punishment other than the focal type under study. This scholarly work concerning the “mixtures” of punishment suggests that punishment should be analyzed as a constitutive package, as the forms of punishment may be interrelated to a certain degree.

While scholarship has shown the dependent nature of these sentences, research has yet to pin down how the presence of one form of punishment could alter the effect of another, which is a plausible scenario given that punishment types are rarely meted out in isolation from one another. Studies to date, to my knowledge, have not studied the packaging – or conditional – effects of various forms of punishment on recidivism. Thus, a central contribution of this dissertation chapter will be to estimate the causal packaging effects of punishment on recidivism, and analyze punishment more closely as it is experienced – as a composite of multiple forms, intensities, and effects of punishment.

Current Chapter

This dissertation chapter uses comprehensive court data for the state of Minnesota from 2004-2015 to investigate the causal impacts of punishment packaging on
recidivism. This chapter will be among the first studies to examine the impacts of multiple different forms and intensities of punishment *simultaneously* on recidivism. Further, it leverages an instrumental variable design that exploits the quasi-random assignment of judges within Minnesota Judicial Districts at the felony-level to estimate the causal effects of the punishment packaging on recidivism.

Secondly, this chapter will examine to what extent sentenced incarceration, probation, and LFO amounts are moderated by the confluence of various forms and intensities of punishment into a packaged sentence. Previous scholarship has also noted heterogeneous treatment effects of incarceration (e.g., Mears et al. 2015), with incarceration’s effects varying by individuals’ “stake in conformity” (Sherman et al. 1992), where negative effects of punishment are found more frequently amongst groups that “buy in” to the conventional social status quo. Additionally, significant scholarship highlights the economic and social strains of punishment and criminal records (e.g., Pager 2003; Western 2002; Stewart and Uggen 2020). Following these findings, this chapter also assesses the extent to which defendant-level disadvantage (measured by the presence of a public defender – see Data and Methods section below), alters the causal impacts of each form of punishment.

I first begin by discussing the court administrative data used in this analysis, and then describe the random judge instrumental variable (IV) design, along with its inherent statistical and design assumptions. Then, I present the results of the IV modeling strategy, and close with a discussion of the findings for the sociological study of punishment and recidivism.
Data and Measures

Minnesota Court Administrative Data

This empirical dissertation chapter draws on an extract of administrative criminal court data from the Minnesota Court Administrator’s Office (SCAO) from 2004-2015. The SCAO data includes resolute charge-, case-, and defendant-level information and events, including defendant demographics. The extract comprises all filed criminal cases in Minnesota from 2004-2015, with sentencing outcome information for all cases resolved before 2019. To identify unique persons in the data, probabilistic deduplication (Forest and Eder 2022) via a partially supervised algorithm\(^1\) with defendant full names and birthdays was used to create unique person identifiers, in order to identify cases within individuals. Legal Financial Obligation (LFO) records come primarily from the Minnesota Court Information System (MNCIS), although Minnesota’s two largest counties by population, Hennepin and Ramsey, used a separate record keeping system, Violations Bureau Electronic System (ViBES) until 2010. MNCIS captures amounts-sentenced, collected, and outstanding- of all three forms of LFOs, and the ViBES data contain case-level information on fines and fees.

I restrict the sample to just felony-level initial cases for three important reasons: 1) felony cases have the least missing data on defendant characteristics due to the in-person nature of felony-level sentencing (see below), 2) I can obtain comprehensive criminal history and offense severity scores at the felony-level, but these are not available for lower charge levels, and 3) felony cases ensure that a judge is present, which is a key to

---

\(^1\) This algorithm calculates record similarity using the affine gap distance on the full name and birthday string fields, then uses human-classified matched outcomes to estimate via iterated regularized logistic regression the probability of a match given the record similarity.
my causal identification strategy. These data are aggregated to an individual-level dataset where each row represents a person’s first instance of sentencing along with relevant charge, case, and defendant characteristics.

Recidivism, the focal outcome in this chapter, is measured as an individual’s *second post-sentence or post-release conviction* as recorded in the SCAO data that contained a conviction of misdemeanor level or higher. Specifically, a person’s second reconviction event is determined as the first subsequent case a) after the sentence date for those not given jail or prison time, or b) after the release date from prison or jail for those sentenced to serve confinement time. This creates equivalent “risk sets” between individuals sentenced to confinement and those not, which removes the effect of incapacitation from the incarceration estimate (Harding et al. 2017). I create follow-up time conditioned event indicators of reconviction across any highest-charge level (excluding petty misdemeanors), which serve as the primary recidivism outcomes in this analysis. I measure recidivism at three different time points (1, 3, and 5 years), creating a unique risk set for each of sentenced individuals that were at-risk for at least the specified time-window, unless having recidivated before the end of the follow-up window. This balances the risk set to be all “surviving” individuals without reconviction up until the follow-up period, as well as individuals reconvicted in that window, effectively excluding individuals that are “right censored” and are not observed or “at risk” for the given length of time (Box-Steffensmeier and Jones 2004). Finally, it should be noted that “re-offense”
not rising to the level of a conviction (e.g., rearrest) or a reconviction in a different state is not captured as a recidivism event.²

The focal punishment measures are also constructed from the SCAO data: incarceration days, probation days, and total LFO order. Incarceration days are calculated by subtracting the number of stayed confinement days (i.e., the days conditionally delayed) and the number of days credited for time served from the ordered number of confinement days. Probation days are simply a measure of the sentenced probation days. LFOs are expressed in terms total dollars sentenced, adjusted for inflation to January 2018 dollars. LFOs in Minnesota (Stewart et al. 2021), and elsewhere (Harris et al. 2017) can take on a variety of forms including direct criminal fines, fees – such as law library fees and mandatory criminal surcharges–, as well as restitution. The ViBES data have incomplete restitution information in Hennepin and Ramsey counties from 2004-2009 (the two largest MN counties by population), so I restrict the LFO measure to include just criminal justice fines and fees, excluding restitution.³ These adjustments ensure the comparability of LFO amounts across space and time. Because the punishment variables are strictly positive and are slightly positively skewed, the punishment variables are expressed using the log(X +1) transformation when included in the multivariate analysis, and therefore the effects of these variables below are interpreted as relative changes.

I also leverage the SCAO data to create a battery of measures of case and defendant characteristics for covariate adjustment. First, I create measures of defendant

² Therefore, recidivism in this study, measured as reconviction, is likely lower than if measured by rearrest amongst comparably situated cases.
³ Amongst first-time felony cases, fines and feed comprise approximately 23.5% of the total ordered LFO amounts in Minnesota counties excluding Hennepin and Ramsey.
race, gender, and age from the SCAO party file. Race and gender data are provided to the court or through a presentence investigation interview and other court documentation devices (e.g., application for public defender). I collapse the SCAO racial categories into six groups: White, Black, Hispanic, Asian/Pacific Islander, Native American, and Other Race (which includes other, multiracial, unavailable, unknown, and refused). Minnesota maintains a state “payables list,” a list of petty misdemeanors and some misdemeanor offenses with defined monetary sanction amounts that are payable without an appearance in court. Therefore, lower offense levels in the SCAO extract have substantial missing data on defendants, whereas higher charge levels have more complete information on defendant characteristics. Given the analysis at hand analyzes first cases with felony convictions, this missingness is substantially reduced. Despite this, I also impute race and gender information within defendants across cases, leveraging information from the defendants’ other cases to impute race and gender in the focal case. Because these measures are taken at each court appearance and/or sentencing, within-person racial and gender self-identification cannot be assumed to be static but instead can change over time (e.g., Liebler et al. 2017) and therefore I impute the modal demographic category when the auxiliary cases are not homogenous across cases within defendants.\(^4\) After this auxiliary information imputation strategy, the demographic data exhibit about 15% and 9% missingness amongst the felony-level first cases for race and gender respectively.

Finally, I create a bevy of other case characteristics from the SCAO data. The presence of a public defender was identified by using the SCAO attorney file, which

\(^4\) In the case of bimodal or multimodal demographic identifications within the auxiliary cases within persons, I randomly sample one of the modes for imputation.
serves as a key proxy variable for defendant disadvantage, as well as their ability to pay in terms of monetary sanctions. Minnesota statute defines public defender eligibility based upon the defendant’s income, liquidity of real estate assets, and other cash convertible assets (MN Stat § 611.17 2021). I also use the charging information within each case to create indicators for whether the case had a violent, drug-related, or alcohol-related conviction, as these have been shown in previous research to associate with LFO amounts (e.g., Harris et al. 2011).

The SCAO data also includes the full name of the sentencing judge, which is used as an instrumental variable in my identification strategy (see details below). Some judges in the SCAO data had very few observed cases, and in the interest of estimation stability I keep cases which had a presiding judge that had at least 5 observed cases over the full study window. After the adjustments for judge case-counts and listwise deletion on missing values of covariates, this results a sample size of 6,909 (1-year), 5,829 (3-year), and 4,654 (5-year) in each respective risk set for analysis.

Minnesota Sentencing Guidelines Data

I merge criminal history and offense severity covariates from the Minnesota Sentencing Guidelines data onto the constructed SCAO data by case number. The Minnesota Sentencing Guidelines Commission keeps detailed data on felony-level sentencing in Minnesota (e.g., Minnesota Sentencing Guidelines Commission 2020), and the presumed sentences in Minnesota are completely dictated by offense severity and criminal history (with separate tables for drug and sex-related cases). The cases’ severity level is determined by only the charges for which a defendant is convicted, and the criminal history score is determined by formulas that account for prior felonies, custody
status, prior gross misdemeanors and misdemeanors, and juvenile adjudications (see Minnesota Sentencing Guidelines Commission 2021). Severity is measured from 1-12 with all the murder severity distinctions binned into the 12th category, and the criminal history points are measures from each defendant’s criminal history, and range from 1-6 in the sample. These measures serve as crucial covariates to the current analysis, as criminal history and offense severity are highly associated with both sentences (Light 2022; Johnson 2006) and recidivism (Harding et al. 2017; Huebner and Berg 2011; Chiricos et al. 2007) and could present threats to the identification strategy’s exclusion restriction (see Analytical Strategy section below). However, a number of first-time cases that included a felony-level highest convicted charge in the SCAO data did not have matching criminal history and severity information in the MNSG data (n=5747). Given the importance of these measures for the empirical design, the final analytical sample consists of first-time felony convictions for which sentencing guidelines information was available. Descriptive statistics for all variables included in the analysis are relayed in Table 1.

5 Criminal history points can be much higher than 6, but because the sample is each individual’s first instance of a felony conviction, these histories only include gross misdemeanor and misdemeanor convictions, which are weighted less than felony-level cases (Minnesota Sentencing Guidelines Commission 2021).
I begin the analysis by describing the relative frequency distribution of the packages of punishment, which represent combinations of the punishment axes of

**Analytical Strategy**

**Descriptive Analyses**

I begin the analysis by describing the relative frequency distribution of the packages of punishment, which represent combinations of the punishment axes of
incarceration, probation, and monetary sanctions. I then describe the focal outcome of interest, recidivism, at each of the three follow-up lengths (1-year, 3-year, 5-year) by plotting the percentage of the sample having recidivated at each follow-up time. I then stratify these recidivism estimates by the package of the punishment levied, to visually examine the relationship between the packages of punishment and recidivism. Each recidivism plot is faceted by the type of highest reconviction charge – any, felony, gross misdemeanor, and misdemeanor. These descriptive tables and figures will give a sense of how punishment is packaged in Minnesota, and how these packages may pattern recidivism.

*Random Judge Assignment*

Using observational data to identify the impact of punishment on recidivism is susceptible to endogeneity, as unobserved attributes of both the case and the defendant may also impact the probability of reconviction. Therefore, this study leverages the random assignment of judges within Minnesota judicial districts to alleviate endogeneity concerns in the estimation of punishment’s effects. Previous scholarship has utilized this method to look at the sentencing impacts on recidivism (Berube and Green 2007; Green and Winnik 2010; Harding et al. 2017) and employment outcomes (Kling 2006; Harding et al. 2018). Despite some technical complexities in random judge designs, the logic is relatively straightforward and intuitive if judge assignment is as-good-as random: if recidivism rates for judge caseloads are practically indistinguishable, then it is unlikely that the differential use of punishment has a causal impact, whereas if they are different these recidivism differences are likely attributable to the differential treatment dosages of punishment (Loeffler and Nagin 2022).
In these quasi-experimental designs, the random assignment of judges serves as an exogenous instrumental variable (IV). The proposed model for the use of random judge assignment in the context of sentencing is displayed in the Directed Acyclic Graph (DAG) in Figure 1. The aim is to estimate the causal effect of the exposure to punishment on recidivism, which is represented by the arrow from punishment, P, to recidivism, R. The variable U is symbolic of all unmeasured confounders for punishment and recidivism. Under the IV assumptions (see details below), the judge, J, is associated with P, affects R only through P, and the association between J and P is unconfounded, conditional on O, the observed covariates. In short, this model isolates variation in sentencing that is plausibly exogenous to recidivism, and given the IV assumptions are met, represents the causal effect of punishment on recidivism.

Figure 1: IV Model of Recidivism using Random Judge Assignment

---

6 The first assumption, the relevance condition, can also be met if J has no direct causal effect on P, but J and P have common causes. This is unlikely to be the case in the case of judges, as they construct, along with sentencing guidelines and other considerations, the punishment meted out by the sentence.
Granted that the assumption for a valid instrument have been met, the intuition here is that random assignment of judges isolates the variation induced exogenously by judge assignment, thereby allowing causal inference of the effect of punishment on recidivism.

*Instrumental Variables Two-Stage Least Squares (2SLS) LPM*

Following previous scholarship (Harding et al. 2017), I model recidivism as a binary outcome using an instrumental variables (IV) linear probability model (LPM), instrumenting the three focal punishment variables, as well as any subsequent interaction terms, with a battery of judge instruments. To guard against the potential violations of the exclusion restriction, the IV models also include effects of the exogenous, pre-sentence characteristics as well as fixed effects for judicial districts and sentence year, as judges are assigned within Minnesota judicial districts. Judicial districts uniquely encompass Minnesota counties, the judicial district fixed effects account for sentencing patterns between districts, as ecological factors such as court caseload pressure and racial and ethnic composition have been shown to impact sentencing above and beyond individual-level factors (Ulmer and Johnson 2004).

I estimate these IV LPM models using two-stage least squares (2SLS) which uses a system of equations, with a first and second stage, to estimate the causal impact of punishment on recidivism. The “first stage” models, which regress the focal punishment treatments on the instruments, are specified as follows:

\[
P_i = \alpha + \pi J_i + \beta X_i + \theta l + \gamma_i + \epsilon_i
\]

where \(P_i\) represents one of the focal punishment treatment variables, \(J_i\) is a matrix of judge dummy instruments, \(X_i\) are the observed case and defendant characteristics. I also
include fixed effects for judicial district \((\theta_i)\) and sentence year \((\gamma_i)\) to adjust for spatiotemporal heterogeneity in sentencing patterns.\(^7\)

The reduced form ("second stage") models are specified as follows:

\[
R_i = \alpha + \pi_2 J_i + \beta X_i + \theta_i + \gamma_i + \epsilon_i
\]

where recidivism \((R_i)\) is specified as a function of judge instruments \((\pi J_i)\), the observed case and defendant characteristics \((X_i)\), as well as the fixed effects for judicial district \((\theta_i)\) and sentence year \((\gamma_i)\). If the assumptions outlined in the next section are met, the causal estimate of the impact of the punishment indicators of each of the three punishment indicators can be obtained via the ratio of the reduced form coefficients over the first-stage coefficients:

\[
\delta_{IV} = \frac{\pi_2}{\pi_1}
\]

where \(\delta_{IV}\) is a vector of estimated causal effects for each punishment treatment. The estimation procedure described here is done in one step, which yield standard errors adjusted for the strength of the relationship between the instruments and the treatment variables.\(^8\) 2SLS estimates, in this study, effectively estimate the resulting movement in recidivism probability as a result of the change in punishment induced by judges (see Huntington-Klein 2021:481). All models are estimated using the ‘ivreg’ command in the R package ‘AER’ (Kleiber and Zerleis 2008).

---

\(^7\) This makes the estimation of judge effects, as well as punishment effects in the reduced form model, to be within-judicial district, which net out differences in the cultural sentencing practices of each judicial district, as well as over time changes in sentencing patterns across Minnesota.

\(^8\) In effect, this adjusts the standard errors for the fact that the treatment regressors in the second stage are estimated, as opposed to observed.
**IV Design Assumptions**

To serve as a valid instrument, judge assignment must meet two primary conditions. First, the relevance condition requires that the instrument must induce a change in the endogenous treatment variable, the sentenced punishment, and that this change will translate into a sufficiently strong relationship between them. This would mean that judges have discretion to determine the sentence, above and beyond other observed case and defendant characteristics. Minnesota represents a unique institutional structure for the judge IV design, as it was the first state to adopt a robust set of sentencing guidelines at the felony-level (Tonry 1991). These guidelines give a presumed sentence based upon criminal history and case severity, and then leave room for judicial discretion within these parameters in sentencing. The relevance assumption can be empirically assessed by examining whether judges impact sentencing independently of other observed covariates.

I examine the relevance of judges in criminal sentencing first by visually depicting the judge effects from the “first-stage” model. Figure 2 visualizes the judge effects of punishment amounts sentenced across all three axes of punishment (incarceration, probation, and LFOs). Each vertical bar represents the effect of an individual judge on each respective form of punishment, after accounting for the observed case and defendant characteristics. The variation across judges within judicial districts is obvious in the plots. However, estimates from an IV design can be susceptible to bias if instruments are weak, meaning they do not impact sentencing to a significant degree.

---

9 This makes each judge value relative to the referent judge.
The classic statistical test for weak instruments is to calculate a joint significance F-test on the inclusion of the instruments above and beyond the other observed covariates. Traditionally, F-statistics above 10 have been used as a benchmark (Stock and Yogo 2005), but in studies with many instruments and a large sample size, F-statistics may not be informative, as the number of instruments and the sample size are included in the calculation of the statistic (Harding et al. 2017). Generally, the F-tests for the specifications in this study do not reach this threshold, which potentially could indicate that judges serve as weak instruments.

Therefore, I calculate Shea’s partial $R^2$ (Shea 1997) for each specification, which is less susceptible to the number of instruments and sample size. Shea’s partial $R^2$ quantifies the amount of variance in the treatments explained by the instruments after accounting for the associations with covariates, which consistently show between approximately 5% and 13% of explained variation in sentences by the judge indicators depending on the type of punishment and specification. These values are comparable in magnitude to previous research using this research design (Harding et al. 2017), and, given the sentencing guidelines in Minnesota at the felony-level, represent substantial

---

10 This occurs because, in real world data samples, even valid instruments will likely be associated with the error term just by chance, which, in the presence of weak instruments, weakens the validity of the estimates.
11 Shea’s partial $R^2$ (Shea 1997) is calculated by 1) calculating the first stage regressions for each endogenous variable and obtain fitted values for each, 2) obtaining the residuals from a regression of the focal independent variable on the other endogenous variables and included covariates (effectively partialling out the effects of the other variables), 3) obtaining the residuals of regressing the fitted values of the focal endogenous regressor on the fitted values for the other endogenous variables and the observed covariates (effectively isolating the part of the endogenous focal variable due to the instrument), and 4) obtaining the $R^2$ value from regressing 2 on 3. This amounts to the proportion of variation accounted for by the judge instruments on the focal endogenous variable, after the effects of the other endogenous variables and exogenous covariates have been removed.
12 The calculated Shea’s partial $R^2$ values for each endogenous variable in each specification are included in the regression tables in the Results section.
movement in sentences as the result of judge discretion above and beyond the observed covariates. Additionally, econometric research has shown that including many weak instruments in an IV specification can result in consistent estimation (Chao and Swanson 2005), and in some cases when some within the battery of weak instruments are invalid (Kolesar et al. 2015). In sum, I argue that the relevance condition is satisfied due to the Shea’s partial R² values and the inclusion of many instruments.

The second primary assumption in IV designs is the exclusion restriction, which requires that a) judges are as good as randomly assigned, and b) the judges only impact recidivism through the causal path through the sentence. In essence, the quasi-randomization ensures that the judge to whom a case is assigned is uncorrelated with any other confounding variable that also influences recidivism, thereby making the causal

---

13 The idea here is that in the first stage, where the instrument effects are combined together, the valid instruments can, in some circumstances, wash out the effects of the invalid ones (Kolesar et al. 2015).
path from judge assignment to recidivism flow through the focal treatment variables, the types of punishment. In other words, the randomization assures that judge assignment is not correlated to the error term of the model, making it an exogenous source of variation able to identify the causal effect of sentences on recidivism. Thus, it is imperative to assess whether the assignment of judges is as good as random within Minnesota Judicial Districts. While Minnesota State Courts do not have a universal, blanket judge assignment policy across judicial districts, I evaluate the “as good as random” status of Minnesota judges by checking for differences in observed case characteristics across judges to test whether judicial district judge assignment has behaved quasi-randomly based upon observed characteristics (Nagin and Snodgrass 2013). To test for quasi-randomization, I test for mean equality across judges within judicial districts across a battery of observed case characteristics using one-way ANOVA analysis with $\alpha=.05$. Results of these tests indicate that six of the ten judicial districts have significant ANOVA tests in three or fewer of the observed covariates. However, given the large sample size, the F-statistics are likely to be statistically significant even in the presence of small amounts of variation, and the magnitudes of the F-tests are generally small considering the sample size. While these tests do indicate some level of judge assignment non-randomness in Minnesota, particularly in certain judicial districts (i.e., District 5), I include numerous case- and defendant-level covariates in all IV specifications, which will adjust for differences across judges in observed defendant and case characteristics. Therefore, the validity of the exclusion restriction, in the following analysis, relies upon the assumption that judges are assigned in a quasi-random nature conditional on the
observed covariates.\textsuperscript{14} I also calculate Wu-Hausman tests for endogeneity, which indicate whether endogeneity exists between punishment and recidivism, which would justify the use of the IV design.\textsuperscript{15}

The exclusion restriction also maintains that the instrument can only affect recidivism through its impact on the sentence. Any other influence of judges must be both directly and indirectly inconsequential for recidivism except for the path through the criminal sentence. Violations to the exclusion restriction here could include whether judges’ courtroom behavior influences the likelihood of recidivism beyond the stated sentence (e.g., admonishments), or if prosecutors are responsive to judge’s by changing plea bargaining or charging decisions – which would contaminate the judge effect with that of a prosecutor effect inducing a potential correlation between judge harshness and recidivism not due to the sentence (Harding et al. 2017). Due to the inclusion of many instruments, I also calculate Sargan overidentification tests for each specification, which tests for the exogeneity of the judge instruments.\textsuperscript{16} Across all model specifications, the Sargan tests are not statistically significant, lending evidence towards the validity of

\begin{footnotesize}
\textsuperscript{14} Assignment based on unobserved case or defendant characteristics would represent a violation of the exclusion restriction, as this would open back door paths between punishment and recidivism. However, given that I cannot verify that judge assignment is as good as random with respect to unobserved characteristics, I argue that selective case assignment, to the extent it occurs, is likely to follow characteristics that are accounted for with the inclusion of the covariates, such as criminal history, case severity, and case types (e.g., drug vs. violent).

\textsuperscript{15} The Wu-Hausman test is calculated by taking the residuals from the first stage regression and including them as a regressor in the second stage. The test statistic is then derived from the T-statistic on the coefficient for the residual regressor. This effectively tests whether the "contaminated" variation, the variation being removed from the instrumented variables in the first stage, affects the outcome and therefore would lend evidence of endogeneity. The null hypothesis is indicative of no significant endogeneity between punishment and recidivism.

\textsuperscript{16} Sargan tests are calculated by first estimating the 2SLS regression, obtaining the residuals, and then regressing the residuals on the instruments. This effectively tests for association between the instruments and the residual variance in recidivism, leveraging the excess information from the overidentified model. The null hypothesis is indicative of instrument exogeneity, and a significant test would indicate that the punishment is not the sole path to the outcome for the effects of judges.
\end{footnotesize}
judges as an instrumental variable for estimating the causal effect of punishment on recidivism.

IV designs estimate the “local average treatment effect” (LATE), which is local in the sense that the effect is estimated for the subset of cases for which the judge induces a change in punishment (Angrist, Imbens, and Rubin 1996). Thus, this IV design estimates the effect for individuals at the margin of multiple different treatment potentialities, as in some cases (e.g., homicide) there likely exists less between-judge variation in sentencing severities. The LATE interpretation in the presence of effect heterogeneity, requires an additional “monotonicity” assumption, which requires that judge effects are consistent in effect across defendants. In other words, “a judge who imposes more punitive sentences than her colleagues to some offenders does not also impose more lenient sentences than her colleagues to others (Harding et al. 2017: Appendix 8).” This is commonly called the “no defiers” condition in the IV literature, in the sense that the effect on those assigned to certain instruments, or levels of instruments, follow a consistent rank ordered pattern. Following Harding et al. (2017), I test for monotonicity by including interaction terms between the judge instruments and a number of key case characteristics (e.g., criminal history, case severity), which allow the judge effects to vary by case and defendant characteristics. Specifications with and without these extra instruments are substantively similar, indicating that monotonicity does not appear to be a concern in these data (at least in terms of the case observables) and therefore I present the more parsimonious IV specifications. I first estimate 2SLS models of each follow-up recidivism event, including each of the focal punishment treatments in the models, allowing for the assessment of the relative causal effects of each punishment form on recidivism. I then present
specifications that include two-way interactions between the forms of punishment, to examine whether the effects of punishment are conditioned by the packaging of punishment forms together. Finally, I present IV models with interaction terms between each form of punishment and the public defender indicator, to assess whether the effects of punishment may be different for more disadvantaged defendants.

Results

Packages of Punishment

Table 2 displays the frequency distribution of the packages - or combinations - of the three different axes of punishment of incarceration, probation, and LFOs in the analytic sample. At the felony-level the “full package” inclusive of all three forms of punishment is the most frequent, comprising approximately 39% of cases. This category is followed by the incarceration LFO and probation LFO packages, each with a relative frequency of ~26% and ~16% respectively, making these packages the most frequent after the “full package.” In comparison, defendants receiving incarceration and probation together without a LFO is relatively uncommon, at just about 7.5% of cases. Defendants experiencing any form of punishment in isolation at the felony-level is also fairly rare, with just a combined ~11% of felony-level defendants experiencing a lone punishment type at sentence. Finally, defendants sentenced to no punishment across the sample is exceptionally rare.\(^{17}\) This corroborates previous scholarship that punishment is experienced as a package, rather than in isolation (e.g., Harris 2016), and this suggests

\(^{17}\) These are primarily cases in which sentenced incarceration time was either a) stayed, or b) effectively zero due to credited days.
that punishments have at least the *opportunity* to interact with each other to pattern outcomes for those subject to its purview.

<table>
<thead>
<tr>
<th>Package</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>30</td>
<td>0.40</td>
</tr>
<tr>
<td>LFO Only</td>
<td>97</td>
<td>1.30</td>
</tr>
<tr>
<td>Probation Only</td>
<td>252</td>
<td>3.38</td>
</tr>
<tr>
<td>Incarceration Only</td>
<td>480</td>
<td>6.44</td>
</tr>
<tr>
<td>Incarceration Probation</td>
<td>557</td>
<td>7.48</td>
</tr>
<tr>
<td>Probation LFO</td>
<td>1175</td>
<td>15.77</td>
</tr>
<tr>
<td>Incarceration LFO</td>
<td>1949</td>
<td>26.16</td>
</tr>
<tr>
<td>Full Package</td>
<td>2909</td>
<td>39.05</td>
</tr>
</tbody>
</table>

*Recidivism*

The focal recidivism indicator, again measured as reconviction in this study, is graphed at each follow-up time (1-year, 3-year, 5-year) in Figure 3. In the first panel, recidivism of cases which included a highest reconvicted charge of any variety shows that as of 1-year, just north of 25% of the sample had been reconvicted, and this number increases over time, with ~45% reconvicted as of 3-years and ~57% reconvicted as of 5-years of follow-up. In addition, the increases in the likelihood of recidivism show greater increases from years 1-3 than from years 3-5, suggesting that the hazard of recidivism increases the most in the initial years post-treatment, followed by slight declines in the time after follow-up year 3. These patterns are consistent with previous recidivism research (e.g., Harding et al. 2017; Durose et al. 2014). The other panels, which depict recidivism at each reconviction level, also show similar over-time patterns, with misdemeanors being the most likely highest reconviction charges, followed by felony and gross misdemeanor charges. Figure 1 indicates that even when measuring recidivism via
reconviction, as opposed to rearrest or arraignment, recidivism amongst those convicted with a first-time felony charge is considerably high, approaching 60% as of 5-years.

Figure 3: Recidivism Events by Highest Reconvicted Charge

Figure 4 displays similar recidivism calculations as Figure 3, but stratifies each recidivism outcome by the punishment package received upon sentencing of the base case. Overall, the recidivism estimates do not vary greatly with different types of punishment packages, suggesting that punishment effects may be weaker in magnitude in patterning recidivism outcomes. In the overall recidivism panel, the “probation only,” “incarceration probation,” and “incarceration only” packages have the highest likelihoods of recidivism over the follow-up periods, all of which end up above 60% recidivism as of year 5. This could be because incarceration and/or probation have criminogenic effects on recidivism, or because these punishments are associated with case or defendant
characteristics that spell higher likelihoods of recidivism (e.g., criminal history).

Punishment packages that include LFO cases are slightly lower in the overall recidivism panel, but all result in at least 50% recidivism by year 5. Generally, the highest charge degree-conditioned panels exhibit similar between-package variation, albeit with the varying baselines of recidivism likelihoods displayed in Figure 3. In sum, these figures suggest that the effects of incarceration are likely to be slight, and that those given probation, incarceration, or a combination therein may have greater likelihoods of recidivism. It should be noted that this visualization merely depicts the bivariate association between package type and recidivism, thereby a) being susceptible to omitted-variables bias, and b) not directly assessing the variation in punishment intensity (e.g., amount of sentenced incarceration). The following instrumental variables models will partition the variance in punishment to remove the threat of omitted variables bias, as well as directly model the intensity dynamics of each mode of punishment.
I now move to presenting the 2SLS instrumental variable model estimates that, granted the assumptions defined in the Analytical Strategy section are met, represent the causal effect of punishment intensities on recidivism. Table 3 presents the primary specifications of recidivism, with each punishment treatment effect representative of the net causal effect of each form of punishment. Modeling reconviction as of 1-year of follow up, none of the punishment measures have statistically significant effects on recidivism. Albeit statistically nonsignificant, the sign of the effects varies across models, with incarceration and LFO amounts having negative sign, and probation having positive sign. These findings suggest that, independently, each form of punishment do not appreciably pattern recidivism in the first year of risk. It should be noted that the race coefficients estimated here primarily reflect system activity, as opposed to individual
behavior (National Academy of Sciences, Engineering, and Medicine 2022), and I discuss this further when interpreting these race coefficients below.

The second model in Table 3 relays a specification that measures recidivism with a 3-year follow-up window. In contrast to the year following the start of the period “at-risk,” the probation effect almost doubles in effect (.008 to .015) and is statistically significant, indicating that a one percent increase in probation days leads to a .00015 (.015*log(1.01) = .00015) increase in the probability of recidivism. Albeit statistically significant, this effect is small in magnitude, but suggests that probation is slightly criminogenic in its impact. Put differently, a 100% increase in probation (i.e., a doubling of a probation length), is related to a .01 increase in the probability of recidivism, showing that long probation sentences, which were plentiful in Minnesota during the observation period (Minnesota Justice Research Center 2019), can have modest criminogenic effects. This effect is also apparent in the 5-year follow-up specification in Table 3 with a one percent increase in recidivism leading to a .00017 increase in the probability of recidivism as of 5-years. Together, these coefficients suggest that probation becomes more criminogenic over time, and longer probation sentences are particularly, although weakly, criminogenic.
In contrast, LFO fine and fee amounts show a statistically significant negative effect on 3-year recidivism with a one percent increase in fine and fee order leading to a 

\((-0.021\times\log(1.01))\) .0002 reduction in the probability of recidivism, net of other variables. Again, this effect is quite weak in magnitude, but cannot be explained away by LFOs being associated with less serious cases, as this is accounted for in the IV design, as well as through the inclusion of offense severity as a control variable. Therefore, this effect is consistent with the notion that cases given higher LFOs have slightly lower probabilities of recidivism. However, this effect is only statistically significant in the 3-year follow up model, which suggests that the effect of LFOs, while negative, are teetering on the verge of statistical significance.

Table 3: IV 2SLS Models of Recidivism

<table>
<thead>
<tr>
<th>Variable</th>
<th>1 Year</th>
<th>Follow-Up Period</th>
<th>5 Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infant (log)</td>
<td>-0.012 (0.016)</td>
<td>-0.013 (0.013)</td>
<td>-0.001 (0.013)</td>
</tr>
<tr>
<td>Prob (log)</td>
<td>0.008 (0.006)</td>
<td>0.015* (0.007)</td>
<td>0.017 (0.007)</td>
</tr>
<tr>
<td>LFO (log)</td>
<td>-0.004 (0.007)</td>
<td>-0.021* (0.009)</td>
<td>-0.018 (0.010)</td>
</tr>
<tr>
<td>Black</td>
<td>0.048** (0.017)</td>
<td>0.067** (0.022)</td>
<td>0.058* (0.025)</td>
</tr>
<tr>
<td>Hisp.</td>
<td>-0.040* (0.020)</td>
<td>-0.078** (0.025)</td>
<td>-0.108*** (0.028)</td>
</tr>
<tr>
<td>Nat. Am.</td>
<td>0.058** (0.018)</td>
<td>0.101*** (0.022)</td>
<td>0.108*** (0.023)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.003 (0.037)</td>
<td>-0.011 (0.047)</td>
<td>0.015 (0.052)</td>
</tr>
<tr>
<td>Other</td>
<td>-0.019 (0.042)</td>
<td>-0.061 (0.053)</td>
<td>-0.123* (0.061)</td>
</tr>
<tr>
<td>Male</td>
<td>0.047** (0.014)</td>
<td>0.068*** (0.017)</td>
<td>0.066*** (0.019)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.006*** (0.0004)</td>
<td>-0.008*** (0.001)</td>
<td>-0.008*** (0.001)</td>
</tr>
<tr>
<td>Pub. Def.</td>
<td>0.068*** (0.013)</td>
<td>0.095*** (0.016)</td>
<td>0.099*** (0.018)</td>
</tr>
<tr>
<td>Violent</td>
<td>0.049*** (0.015)</td>
<td>0.044* (0.018)</td>
<td>0.058** (0.020)</td>
</tr>
<tr>
<td>Drug</td>
<td>0.026 (0.046)</td>
<td>0.018 (0.029)</td>
<td>0.032 (0.023)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>0.063 (0.024)</td>
<td>-0.015 (0.029)</td>
<td>0.040 (0.032)</td>
</tr>
<tr>
<td>Crm. Hist.</td>
<td>0.010** (0.005)</td>
<td>0.020** (0.006)</td>
<td>0.021** (0.007)</td>
</tr>
<tr>
<td>Severity</td>
<td>-0.007 (0.004)</td>
<td>-0.012* (0.005)</td>
<td>-0.015* (0.006)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.520*** (0.129)</td>
<td>0.610 (0.341)</td>
<td>0.577 (0.341)</td>
</tr>
</tbody>
</table>

| District FE | Yes | Yes | Yes |
| Sentence Year FE | Yes | Yes | Yes |
| IV F (Incar.) | 2.172*** | 2.348*** | 2.554*** |
| IV F (Prob.) | 3.254*** | 3.693*** | 4.762*** |
| IV F (LFO) | 2.630*** | 2.537*** | 2.903*** |
| Wu-Hausman | 0.967 | 3.860** | 2.876* |
| Sargan | 276.020 | 230.279 | 178.988 |
| Shea R² (Incar.) | 0.073 | 0.077 | 0.077 |
| Shea R² (Prob.) | 0.163 | 0.114 | 0.129 |
| Shea R² (LFO) | 0.095 | 0.095 | 0.084 |
| Observations | 5,999 | 5,829 | 4,654 |

* p<0.05; ** p<0.01; *** p<0.001
IV: Assigned Judge
Besides punishment, a number of other variables significantly pattern recidivism. First, Black defendants are associated with higher likelihoods of recidivism across all follow-up periods (1yr = .048, 3yr = .067, 5yr = .058, p<.05), which is consistent with previous recidivism research (e.g., Chiricos et al. 2017). However, this coefficient, as well as the other racial coefficients as a whole, bear interpretive caution: wrapped up in the effect of defendant race is the associated criminogenic effects of racial segregation (Peterson and Krivo 2010; Mears et al. 2008; Massey and Denton 1993), racial inequalities and associated strains (Kirk 2020; Kirk 2009; Reisig et al. 2007), racialized policing practices (Knox et al. 2020; Goncalves and Mello 2021; Atiba Goff 2012; Beckett et al. 2006; Gelman et al. 2007; Weitzer and Tuch 2005), and broader structural racism and mass criminalization (Hinton and Cook 2021; Alexander 2012) to which Black defendants are subject. In other words, wrapped up in this ostensibly simple coefficient is not just a difference in racial identification, but a plethora of associated social inequities, relationships to social institutions, and social interactional processes that play out in the time period post-sentencing. Similarly, Native American defendants are associated with higher recidivism probabilities, and, in contrast, Hispanic defendants are associated with lower probabilities of recidivism (1-year = -.040, 3-year = -.078, 5-year = -.108, p<.05), both of which are patterns consistent with previous research (Chiricos et al. 2007; Kubrin and Stewart 2006; Vigessa 2013). Additionally, the demographic variables of age and gender are significantly associated with recidivism across the follow-up specifications, with older defendants having lower-likelihoods of recidivism. Both the age (Uggen 2000) and gender (Huebner and Pleggenkuhle 2013) patterns are also common in recidivism research (see also Piquero et al. 2015). The indicator for a violent
conviction was also significantly associated to increased probabilities of recidivism across the models, as was criminal history, with a one percent increase in criminal history associated with .00016, .00020, and .00021 increases in recidivism probability across the three 1-year, 3-year, and 5-year follow-up periods respectively.

Overall, the IV results in Table 3 are suggestive of a weak, albeit statistically significant, criminogenic effect of probation, that modestly increases over the follow-up periods. In contrast to previous research (Harding et al. 2017), incarceration is not shown to significantly pattern recidivism, suggesting that increasing or decreasing incarceration intensity (as opposed to its binary application) is unlikely to impact recidivism. Finally, LFO fines and fees appear to exhibit a negative, albeit weak, causal relationship to recidivism, but the null hypothesis for the LFO effect is only rejected in the 3-year follow-up specification. Therefore, caution is warranted in interpreting this effect, as it is not robust and on the verge of statistical nonsignificance. On the whole, these empirical results are suggestive that the impact of punishment intensities on reconviction amongst those sentenced is relatively weak.

The following regression tables present the first models that examine how the punishment intensities may condition the impact of one another, which begins to analyze some of the potential packaging dynamics of punishment, and how they may interact to shape recidivism. Table 4 includes identical specifications to those depicted in Table 3, with the addition of an interaction term between logged incarceration and logged probation. Across all three follow-up windows, the interaction term is statistically nonsignificant and close to 0, indicating that the modest criminogenic effect of probation, as evidenced in Table 3, is not moderated by the amount of incarceration given. These
findings suggest that probation, regardless of the amount of incarceration meted out, has a consistent positive and criminogenic influence on recidivism likelihoods, potentially indicative of a “surveillance” effect, where probation “picks up” future instances of criminal behavior more frequently, or this could alternatively be due to the increased criminalization of behaviors that garner probation violations.

<table>
<thead>
<tr>
<th>Table 4: IV 2SLS Models of Recidivism - Incar. X Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recidivism</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>Incarceration(log)</td>
</tr>
<tr>
<td>Probation(log)</td>
</tr>
<tr>
<td>LFO(log)</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Nat. Am.</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Pub. Def.</td>
</tr>
<tr>
<td>Violent</td>
</tr>
<tr>
<td>Drug</td>
</tr>
<tr>
<td>Alcohol</td>
</tr>
<tr>
<td>Crim. Hist.</td>
</tr>
<tr>
<td>Severity</td>
</tr>
<tr>
<td>IncarXProb</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>

**Notes:** *p<0.05; **p<0.01; ***p<0.001
IV: Assigned Judge

I start to see some evidence towards the packaging effects of punishment in Table 5, which relays specifications similar to those in Table 3, but allowing the effects of probation and LFOs to be conditional upon one another. The 1-year follow up
specification shows a statistically significant interaction positive interaction effect, with higher LFO amounts moderating the causal effect of probation to become more criminogenic. Specifically, a one percent increase in LFO amount increases the effect of probation on recidivism by .008 (p<.05). The main effects for both probation and LFOs are statistically significant and negative, suggesting that each form punishment, when coupled with low amounts of the other, is causally linked to lower likelihoods of recidivism. Taking the main effects and the interaction effect into account together, these coefficients are suggestive that these punishment effects, if experienced in isolation, exhibit a suppressing effect on recidivism, but each becomes more positive as more of the other form of punishment is added to the sentence. Figure 5 displays the dynamics of these causal effects, where the probation effect at lower levels of sentenced LFOs is negative, and as LFO amounts increase, so does the criminogenic direction of probation.

While this interaction effect is not present in the other follow-up periods, this could suggest that the packaging effect of probation and LFOs are particularly harmful in the first year of risk, where LFO payment concerns, coupled with the economic and social strain associated with the presence of a felony-level criminal record (e.g., Pager 2003), are at their height. It should also be noted that a low frequency of felony-level defendants experience either of these forms of punishment in isolation, suggesting that, in the majority of cases, the influences of these forms of punishment are likely to be null or positive as they are frequently packaged together (see also Table 2).
Table 5: IV 2SLS Models of Recidivism - Prob. X LFO

<table>
<thead>
<tr>
<th></th>
<th>1 Year</th>
<th>Follow-Up Period</th>
<th>5 Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incar (log)</td>
<td>-0.011 (0.019)</td>
<td>-0.013 (0.012)</td>
<td>-0.0003 (0.013)</td>
</tr>
<tr>
<td>Prob. (log)</td>
<td>-0.028* (0.013)</td>
<td>0.001 (0.015)</td>
<td>0.026 (0.016)</td>
</tr>
<tr>
<td>LFO (log)</td>
<td>-0.043** (0.014)</td>
<td>-0.037* (0.017)</td>
<td>-0.007 (0.018)</td>
</tr>
<tr>
<td>Black</td>
<td>0.049** (0.018)</td>
<td>0.067** (0.022)</td>
<td>0.055* (0.025)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.039 (0.021)</td>
<td>-0.077** (0.025)</td>
<td>-0.109*** (0.028)</td>
</tr>
<tr>
<td>Nat. Am.</td>
<td>0.055** (0.018)</td>
<td>0.100*** (0.022)</td>
<td>0.109*** (0.023)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.009 (0.038)</td>
<td>-0.009 (0.047)</td>
<td>0.014 (0.052)</td>
</tr>
<tr>
<td>Other</td>
<td>-0.027 (0.043)</td>
<td>-0.066 (0.053)</td>
<td>-0.121* (0.061)</td>
</tr>
<tr>
<td>Male</td>
<td>0.042** (0.014)</td>
<td>0.066** (0.017)</td>
<td>0.067*** (0.019)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.006*** (0.005)</td>
<td>-0.008*** (0.001)</td>
<td>-0.008*** (0.001)</td>
</tr>
<tr>
<td>Pub. Def.</td>
<td>0.068** (0.014)</td>
<td>0.096*** (0.016)</td>
<td>0.099*** (0.018)</td>
</tr>
<tr>
<td>Violent</td>
<td>0.051*** (0.015)</td>
<td>0.047* (0.018)</td>
<td>0.050** (0.020)</td>
</tr>
<tr>
<td>Drug</td>
<td>0.023 (0.016)</td>
<td>0.017 (0.020)</td>
<td>0.032 (0.023)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>0.001 (0.024)</td>
<td>-0.014 (0.029)</td>
<td>0.036 (0.032)</td>
</tr>
<tr>
<td>Crim. Hist.</td>
<td>0.008 (0.006)</td>
<td>0.017* (0.007)</td>
<td>0.023** (0.008)</td>
</tr>
<tr>
<td>Severity</td>
<td>-0.009* (0.004)</td>
<td>-0.012*** (0.005)</td>
<td>-0.014*** (0.005)</td>
</tr>
<tr>
<td>Prob X LFO</td>
<td>0.008** (0.002)</td>
<td>0.003 (0.003)</td>
<td>-0.002 (0.003)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.722*** (0.145)</td>
<td>0.676 (0.348)</td>
<td>0.530 (0.346)</td>
</tr>
</tbody>
</table>


Note: *p<0.05; **p<0.01; ***p<0.001
IV: Assigned Judge

Figure 5: Probation X LFO Interaction Plot

1-Year Recidivism IV Model

LFO (log)
-2 SD    -1 SD    Mean   +1 SD    +2 SD

0.8
0.6
0.4
0.2
0.0

Pr(Recidivism)

Probation Days (log)
0
5
10
15
20
25
30
Table 6 contains similar IV specifications to those discussed in the previous two tables, but instead includes the final two-way interaction between the three forms of punishment: incarceration days and LFOs. Across all three follow-up windows, the interaction effect between incarceration and LFOs is statistically nonsignificant and close to zero. This indicates that the effects of incarceration or LFOs are not conditional upon one another. In sum, these IV models presented thus far show that overall, punishment patterns have weak, if any, causal impacts on recidivism, but that the impositions of hefty LFOs can exacerbate the criminogenic effect of probation. I now turn to a final set of interaction models that test for heterogeneity in the estimated causal effects across different levels of defendant disadvantage.

Table 6: IV 2SLS Models of Recidivism - Incar. X LFO

<table>
<thead>
<tr>
<th>Recidivism</th>
<th>1 Year</th>
<th>3 Year</th>
<th>5 Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incar. (log)</td>
<td>0.007 (0.019)</td>
<td>-0.007 (0.022)</td>
<td>-0.006 (0.027)</td>
</tr>
<tr>
<td>Prob. (log)</td>
<td>0.015 (0.017)</td>
<td>-0.015 (0.029)</td>
<td>-0.023 (0.023)</td>
</tr>
<tr>
<td>LFO (log)</td>
<td>0.009 (0.006)</td>
<td>0.015* (0.007)</td>
<td>0.017 (0.007)</td>
</tr>
<tr>
<td>Black</td>
<td>0.051** (0.018)</td>
<td>0.008** (0.022)</td>
<td>0.007 (0.023)</td>
</tr>
<tr>
<td>Hisp.</td>
<td>-0.040* (0.020)</td>
<td>-0.078** (0.025)</td>
<td>-0.107** (0.028)</td>
</tr>
<tr>
<td>Nat. An.</td>
<td>0.066*** (0.018)</td>
<td>0.101*** (0.022)</td>
<td>0.107*** (0.023)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.092 (0.037)</td>
<td>-0.012 (0.047)</td>
<td>0.016 (0.052)</td>
</tr>
<tr>
<td>Other</td>
<td>-0.015 (0.012)</td>
<td>-0.009 (0.015)</td>
<td>-0.015 (0.014)</td>
</tr>
<tr>
<td>Male</td>
<td>0.045** (0.014)</td>
<td>0.008** (0.017)</td>
<td>0.066** (0.019)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.006*** (0.0005)</td>
<td>-0.008*** (0.001)</td>
<td>-0.008*** (0.001)</td>
</tr>
<tr>
<td>Pub. Def.</td>
<td>0.067*** (0.013)</td>
<td>0.095*** (0.016)</td>
<td>0.099*** (0.018)</td>
</tr>
<tr>
<td>Violent</td>
<td>0.049*** (0.015)</td>
<td>0.044** (0.018)</td>
<td>0.058** (0.020)</td>
</tr>
<tr>
<td>Drug</td>
<td>0.024 (0.016)</td>
<td>0.018 (0.029)</td>
<td>0.033 (0.033)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>0.068 (0.094)</td>
<td>-0.013 (0.029)</td>
<td>0.029 (0.033)</td>
</tr>
<tr>
<td>Crim. Hist.</td>
<td>0.015*** (0.005)</td>
<td>0.020*** (0.006)</td>
<td>0.021** (0.007)</td>
</tr>
<tr>
<td>Severity</td>
<td>-0.007 (0.004)</td>
<td>-0.012* (0.005)</td>
<td>-0.015* (0.005)</td>
</tr>
<tr>
<td>Incar X LFO</td>
<td>-0.004 (0.004)</td>
<td>-0.001 (0.004)</td>
<td>0.001 (0.005)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.452 (0.418)</td>
<td>0.583 (0.352)</td>
<td>0.683 (0.357)</td>
</tr>
</tbody>
</table>

| District FE | Yes | Yes | Yes |
| Sentence Year FE | Yes | Yes | Yes |
| IV F (Incar.) | 2.172*** | 2.348*** | 2.551*** |
| IV F (Prob.) | 3.254*** | 3.693*** | 4.762*** |
| IV F (LFO) | 2.656*** | 2.547*** | 2.303*** |
| IV F (Incar X LFO) | 2.127*** | 2.068*** | 2.305*** |
| Wu-Hausman | 0.627 | 3.860** | 2.187 |
| Sargan | 275.525 | 230.279 | 178.726 |
| Shea R² (Incar.) | 0.057 | 0.06 | 0.05 |
| Shea R² (Prob.) | 0.103 | 0.114 | 0.127 |
| Shea R² (LFO) | 0.058 | 0.064 | 0.054 |
| Shea R² (Incar X LFO) | 0.053 | 0.054 | 0.036 |
| Observations | 6,909 | 5,820 | 4,654 |

Note: *p<0.05; **p<0.01; ***p<0.001
IV: Assigned Judge
Table 7 presents tests of whether the impacts of punishment on recidivism (1-year follow up) vary by the presence of a public defender. While an imperfect proxy of defendant socioeconomic status (as variation in socioeconomic status still ostensibly vary within those who qualify for public defenders), I interact the instrumented punishment measures with the public defender indicator to test whether effects of punishment are more harmful for disadvantaged defendants. The interaction term between logged incarceration days and the public defender indicator is statistically significant and positive, with the presence of a public defender increasing the causal effect of incarceration by .054 (p<.05). Specifically, the main effect of incarceration, indicative of the estimated effect for individuals without a public defender, is statistically significant and negative, with a one percent increase in sentenced incarceration days associated with a .0005 decrease (p<.05) in the probability of recidivism as of 1-year of follow-up. This is an effect of small magnitude, and similarly implies that a doubling of incarceration days reduces the likelihood of recidivism by just ~.04 (-.056*\log(2.00)). This is indicative of very weak returns of reductions of recidivism for the number of incarceration days required to produce it amongst more advantaged defendants.

While incarceration does have a modest negative effect amongst more advantaged defendants, the effect is effectively 0 amongst more disadvantaged defendants with public defenders, as the interaction term washes out the negative main effect. A one percent increase in incarceration days is linked with a -.00002 decrease in the probability of recidivism amongst less advantaged defendant with public defenders, and a general linear hypothesis test of the combined main and interaction effects is not statistically significantly different from 0 (p>.05). This interaction effect is plotted in Figure 6, which
shows the statistically significant negative effect of incarceration for more advantage defendants who can (or are deemed able) to afford an attorney, and the statistically nonsignificant null effect of incarceration amongst more disadvantaged defendants with public defenders. In sum, these interaction models show that incarceration has a weak, albeit statistically significant, negative causal effect on recidivism amongst more advantaged defendants, but the other forms of punishment have causal effects consistent across levels of defendant disadvantage. This suggests that incarceration only exerts an influence, a weak one at that, amongst those who are less likely to end up within its punitive reach (Wacquant, 2009).

Table 7: IV 2SLS Models of Recidivism - Pub. Def. Interactions

<table>
<thead>
<tr>
<th>Recidivism</th>
<th>IncarXPD</th>
<th>ProbXPD</th>
<th>LPOXPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incar (log)</td>
<td>$-0.056^{*}$ (0.024)</td>
<td>$-0.012$ (0.010)</td>
<td>$-0.011$ (0.019)</td>
</tr>
<tr>
<td>Prob (log)</td>
<td>$-0.156$ (0.114)</td>
<td>$0.208^{*}$ (0.083)</td>
<td>$-0.067$ (0.125)</td>
</tr>
<tr>
<td>LFO (log)</td>
<td>$0.006$ (0.060)</td>
<td>$0.030^{*}$ (0.014)</td>
<td>$0.008$ (0.060)</td>
</tr>
<tr>
<td>Black</td>
<td>$-0.003$ (0.007)</td>
<td>$-0.004$ (0.007)</td>
<td>$-0.026$ (0.022)</td>
</tr>
<tr>
<td>Hisp.</td>
<td>$0.160^{**}$ (0.018)</td>
<td>$0.053^{**}$ (0.018)</td>
<td>$0.048^{**}$ (0.017)</td>
</tr>
<tr>
<td>Nat. Am.</td>
<td>$-0.041^{*}$ (0.020)</td>
<td>$-0.041^{*}$ (0.020)</td>
<td>$-0.045^{*}$ (0.021)</td>
</tr>
<tr>
<td>Asian</td>
<td>$0.056^{**}$ (0.018)</td>
<td>$0.037^{**}$ (0.018)</td>
<td>$0.037^{**}$ (0.018)</td>
</tr>
<tr>
<td>Other</td>
<td>$0.001$ (0.007)</td>
<td>$0.008$ (0.007)</td>
<td>$-0.003$ (0.037)</td>
</tr>
<tr>
<td>Male</td>
<td>$-0.019$ (0.030)</td>
<td>$-0.013$ (0.030)</td>
<td>$-0.022$ (0.042)</td>
</tr>
<tr>
<td>Age</td>
<td>$0.045^{**}$ (0.014)</td>
<td>$0.045^{**}$ (0.014)</td>
<td>$0.046^{**}$ (0.014)</td>
</tr>
<tr>
<td>Pub. Def.</td>
<td>$-0.006^{***}$ (0.005)</td>
<td>$-0.007^{***}$ (0.005)</td>
<td>$-0.006^{***}$ (0.005)</td>
</tr>
<tr>
<td>Violent</td>
<td>$0.039^{***}$ (0.015)</td>
<td>$0.034^{**}$ (0.015)</td>
<td>$0.034^{**}$ (0.015)</td>
</tr>
<tr>
<td>Drug</td>
<td>$0.026$ (0.016)</td>
<td>$0.025$ (0.016)</td>
<td>$0.026$ (0.016)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>$0.000$ (0.024)</td>
<td>$0.0004$ (0.024)</td>
<td>$0.006$ (0.024)</td>
</tr>
<tr>
<td>Crim. Hist.</td>
<td>$0.016^{**}$ (0.005)</td>
<td>$0.016^{**}$ (0.005)</td>
<td>$0.015^{**}$ (0.005)</td>
</tr>
<tr>
<td>Severity</td>
<td>$-0.005$ (0.004)</td>
<td>$-0.006$ (0.004)</td>
<td>$-0.007$ (0.004)</td>
</tr>
<tr>
<td>IncarXPD</td>
<td>$0.054^{*}$ (0.027)</td>
<td>$-0.028$ (0.016)</td>
<td>$0.027$ (0.024)</td>
</tr>
<tr>
<td>ProbXPD</td>
<td>$0.039^{*}$ (0.150)</td>
<td>$0.390^{**}$ (0.150)</td>
<td>$0.635^{***}$ (0.167)</td>
</tr>
<tr>
<td>Constant</td>
<td>$0.094^{***}$ (0.157)</td>
<td>$0.390^{**}$ (0.150)</td>
<td>$0.635^{***}$ (0.167)</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01; ***p<0.001
IV: Assigned Judge
Discussion and Conclusion

This empirical dissertation chapter set out to estimate the causal effects of punishment packaging on recidivism, thereby rigorously evaluating the independent impact of each form of punishment and its relationship to recidivism, as well as the conditional, interactive effects between the modalities of punishment. This research not only has important policy implications for how punishment is used by courts, but also evaluates the potential iatrogenic effects of punishment on crime directly at the individual-level. Given the weak effects of punishment here, combined with the harmful economic and social effects of large-scale punishment (e.g., Uggen and Wakefield 2010; Larson et al. 2021; O’Neill et al. 2021), these findings hold sociological implications for the place of punishment as a social institution. These findings are in concert within a growing body of scholarship in the sociology of punishment, in which punishment regimes are not defined by their ability to reduce recidivism or increase public safety, but
are better conceptualized as a broader network of raced and classed social control mechanisms (Garland 2001; Wacquant 2009; Alexander 2012), that do more to (re)produce inequalities than to combat crime.

First, this chapter shows that, in general, punishment at the individual-level, as it is currently practiced, does little to pattern subsequent crime, as indicated by the weak magnitudes of the effects of punishment on crime. These findings suggest that the imposition of greater punitive sanctions do little to impact the likelihood of recidivism. Further, the primary IV specifications yielded longer-term criminogenic effects of probation, suggesting that probation, and the constraints and strains its places on the lives of those sentenced (Phelps 2013) may be a particular mechanism by which the “revolving door” punishment operates. This suggests that probation could act as a surveillance mechanism that makes reconviction of future criminal behavior more likely. Another mechanism could be the increased criminalization of behaviors that make reconviction more likely due to violation of stringent probation requirements.\(^{18}\)

The findings here are generally weaker than found in previous quasi-experimental studies (e.g., Harding et al. 2017), but this could be due to the differences in the comparison groups across studies. In other words, the weak effects of punishment estimated here are quantifying punishment \textit{intensity}, thereby comparing those sentenced to, for example, greater amounts of probation to those given less, as opposed to other studies comparing those with incarceration to those not given incarceration as a whole. This is not necessarily a limitation to the current study, but rather estimates a different

\(^{18}\) The SCAO data does not include information on whether a case was the result of a probation violation or a new conviction.
effect: that of adding more punishment amongst those similarly punished. Another possibility is that previous studies have largely neglected to examine the interrelationship between the three axes of punishment, and therefore accounting for these relationships may have resulted in weaker, albeit less biased, estimates of the effects of punishment. In all, these findings provide evidence of a weak criminogenic effect of probation on recidivism *above and beyond* the impacts of both incarceration and financial penalties. Coupling these findings with the vast research that highlights the other destructive effects punishment can have on the social life of those within its reach (e.g., Uggen and Stewart 2015), the main effect findings suggest that lowering the sentenced amounts of incarceration, probation, and monetary sanctions would likely not harm public safety in terms of recidivism\(^{19}\), while simultaneously reducing the harms induced by punitive practices.

This study further examined the packaging effects of crime, finding that the imposition of LFOs exacerbated the deleterious effects of probation. These findings highlight that while each form of these non-custodial punishments *in isolation* are rather innocuous or recidivism reducing, when packaged together they amalgamate to have pernicious effects on subsequent reconviction. Specifically, high levels of LFO imposition causes the effect of probation to become criminogenic, with higher likelihoods of recidivism amongst those given higher amounts of probation supervision and hefty LFO orders. This identifies a packaging effect of punishment, in that the imposition of a

---

\(^{19}\) The causal effects from the IV models in this analysis, as discussed in the Analytical Strategy section, are local average treatment effects, and therefore estimating the effect for movements in punishment as a result of the assigned judge. Therefore, the effects estimated here are informative regarding changes on the scale within the variation of sentences that are available to judges. The results here are less informative of more radical policy directions, such as large scale decarceration, which ostensibly would involve cases that do not lie within the range of everyday judicial discretion.
one type of punishment influences the magnitude, and the direction, of the effect of another on recidivism. This marks a key finding in this study, that suggests that the imposition of monetary sanctions on top of already arduous probation requirements may be a criminogenic recipe for elevated recidivism likelihoods.

This chapter also examined the interactive effects of punishment and individual-level defendant disadvantage. The causal effect of incarceration was shown to lower the likelihood of reconviction for more advantaged defendants, and was statistically null for more disadvantaged defendants. While the results of these interaction effects are imperfect due the use of a proxy variable, they are suggestive that incarceration’s effect on recidivism is negative for more advantaged defendants, which represents results consistent with the deterrent or rehabilitative effects of punishment (Nagin 2013). These interactive patterns could be due to the increased amount of strain experienced by disadvantaged individuals for comparable incarceration sentences, as this could “wash out” any potential benefits from the experience of incarceration. Another possibility is that the presence of a public defender is picking up differences in the “stakes of conformity” (Sherman et al. 1992) between defendants, and therefore incarceration may only exert its impact on those with stock in the dominant conventional social order and norms. Previous research also indicates that defendants with public defenders often can’t afford bail and fair worse in the trial and plea-bargaining stages (Page et al. 2019; Gupta et al. 2016; Frazier et al. 2022), which could bring longer spells of pre-trial detention and more severe collateral consequences, making the experience of incarceration worse for indigent defendants. This weak negative effect of incarceration is concentrated amongst groups that make up a minority of defendants in the criminal legal system (Wacquant
2009), and therefore the returns to incarceration are concentrated in a minority subset of those punished (~23% of cases did not have public defenders in the analysis sample). Coupled with the weak magnitude of the effect, increases in incarceration are unlikely to noticeable with respect to recidivism to a meaningful degree. The causal effects of both probation and LFOs were not moderated by defendant disadvantage, which is contrary to hypotheses that defendant “ability to pay” may exacerbate the criminogenic effects of punishment (Harris 2016).

Despite providing rigorous quantitative evidence about the packaging effects of punishment, this study also has a few potential limitations of note. First, while care was taken to impute missing defendant demographic information, imputation always carries the risk of biasing results if the true, unobserved value does not match that which is imputed. However, the imputation strategy used in this analysis relied upon relatively few assumptions, relying exclusively on other cases within defendants at different points in time, and the results of the demographic covariates are consistent in direction to that of previous recidivism studies. Further, while the random judge IV design estimates the causal effect of punishment on recidivism granted that the assumptions are met, there is always a possibility that violations to these assumptions exist. I have done work to provide evidence that the relevance condition has been satisfied, as even in the presence of rigid sentencing guidelines at the felony-level in Minnesota, the Shea’s partial $R^2$ values reach thresholds consistent with previous scholarship (Harding et al. 2017), but are still only explaining less than 13% of the residual variation in punishments at best. While this is consistent with conventions in previous work, the sentencing guidelines limit the available discretion of judges, making judges only modestly relevant for criminal
sentences at the felony-level in Minnesota. Additionally, non-randomness was detected to a certain degree in Minnesota judge assignment, which could result in biased IV specifications, which would open up back door paths between the judge instrument and recidivism. However, I adjusted for the observed differences in assignment across within judicial districts, and Sargan overidentification tests are not indicative of violations to the exclusion restriction.

Minnesota is a state characterized by the imposition of comparably punitive probation sentences (Phelps 2017; Minnesota Justice Research Center 2019), but is comparably less punitive than other states in terms of the imposition of monetary sanctions (Harris et al. 2017) and incarceration (Carson 2020). Minnesota is also somewhat unique in terms of the presence of sentencing guidelines at the felony-level. Therefore, the extent to which these packaging effects may generalize to other jurisdictions with different institutional structures and cultures of sentencing is an open question. Finally, recidivism does not solely measure changes in criminal behavior after treatment. While the use of reconviction represents a “high bar” for recidivism in comparison to rearrest or rearraignment, these measures pick up a mixture of alterations in criminal behavior alongside changes in system surveillance and involvement (National Academy of Science, Engineering, and Medicine 2022). Recidivism measures based on self-reports may be more reflective of changes in crime as the result of treatment as opposed to changes in system behavior. While not yet practiced, these new metrics could be used in a national data collection effort that would enhance recidivism research as a whole (National Academy of Science, Engineering, and Medicine 2022).
With the above caveats, this empirical dissertation chapter provides rigorous quantitative evidence of the impacts of punishment packaging on recidivism, finding weak effects across the axes of punishment, and a significant interactive relationship between probation and monetary sanctions. This finding demarcates the confluence of probation and LFOs as a particularly criminogenic influence. This is an area that policymakers, such as the Minnesota Sentencing Guidelines Commission, could explore to reduce recidivism amongst felony-level defendants, although the magnitudes of the criminogenic effects in this study are weak at best. Overall, these findings suggest that punishment at the individual-level does little amongst those subject to its punitive grasp to enhance public safety, and that the packaging of colloquially “less serious” forms of punishment can have detrimental effects on crime post-sentence. The findings here are in line with recommendations made in a recent National Academies of Sciences, Engineering, and Medicine report (2022) that defining “success” post-punishment should not be limited to just recidivism. The current study suggests that recidivism metrics are unlikely to capture much of the total impacts of incarceration, probation, and monetary sanctions in the lives of those punished, as punishment has weak, and sometimes criminogenic, effects. These findings also add to the scholarship documenting the lack of consistent deterrent or rehabilitative capacity of the criminal legal system (e.g., Mears et al. 2015). While punishment largely fails at the individual-level to produce results concurrent with its manifest goals (e.g., future crime prevention), it simultaneously serves as an institution that deepens existing inequalities along the lines of class and race (e.g., Page and Soss 2021; Uggen and Wakefield 2010; Pager 2003; O’Neill et al. 2021; Larson et al. 2021; Wacquant 2009). The results of this chapter buttress the conceptualizations of
the criminal legal system operating as a social institution that does little to improve public safety, and instead contributes to the structural disadvantage of race-class subjugated groups that are characterized by “insufficient governmental attention and too much government oversight, interference, and predation (Soss and Weaver 2017:2).” In conclusion, this chapter suggests that the effects of punishment on crime at the individual-level are relatively weak, and the iatrogenic impacts of punishment are concentrated among the packaging of community supervision and monetary sanctions – marking the amalgamation of less “serious” forms of punishment as criminogenic in their own right. I now turn to the second empirical chapter that brings the analysis of punishment’s effects to the community-level and analyzes how punishment may impact crime through the patterning of neighborhood structure.
Chapter 2: Punishment, Neighborhood Structure, and Crime

Overview

A long tradition of sociological and criminological research has documented the ecological effects of neighborhoods on crime rates. Theories and research on neighborhood ecology highlight the causal importance of neighborhood structure for crime, and, in particular, concentrated disadvantage. However, this literature largely neglects what role punishment, in terms of incarceration, probation, and monetary sanctions, plays in this theoretical model. Building upon previous scholarship that suggests that punishment plays a key role in the (re)production of social inequality and disadvantage, this chapter uses a unique panel dataset -- merging hospital discharge, state court punishment records, socioeconomic and demographic Census data -- to examine how neighborhood punishment loads may a) directly impact rates of violent injury, and b) impact violence indirectly through the patterning of concentrated disadvantage. Results from ZCTA (Zip Code Tabulation Area)-year fixed effect panel models highlight that while incarceration has a direct negative effect on violence, and probation a positive association, both incarceration and probation are positively associated with concentrated disadvantage, which is, in turn, positively and strongly associated with violence. These results reveal the bifurcating effects of punishment on violence at the community level and highlight a criminogenic path of punishment on violence via the augmentation of disadvantage.
Introduction

A long tradition in sociological criminology has placed the etiological focus of crime at the levels of neighborhoods or communities. In contrast to individual-level theories such as deterrence and differential association, neighborhood ecology theories, as well as empirical work in the associated neighborhood effects tradition (Sampson 2013), explicate how informal social controls are generated, patterned, and impact the ecological space in which individuals are embedded. In the words of Sampson (2013), “neighborhood contexts are important determinants of the quantity and quality of human behavior in their own right (5).” These theories have their roots in the ecological tradition of the Chicago school (Park and Burgess 1925), and, in essence, maintain that neighborhood characteristics do not directly cause crime, but these characteristics are only important for crime insofar as they impact the mediating force of informal social control (see Bursik 1988) and collective efficacy (Sampson et al 1997). This theoretical tradition ties into classic Durkheimian notions of social solidarity and crime (e.g., Durkheim 1893), where social solidarity, as opposed to punishment, is the key driver in variations of crime, and neighborhood contexts can shape the vigor, direction, and efficacies of social cohesion.

The neighborhood effects tradition places the causal locus on the ecological characteristics of neighborhoods, and how these characteristics shape neighborhood crime through the intervening mechanisms of social cohesion (Sampson and Groves 1987), collective efficacy (Sampson et. al. 1997), and legal cynicism (Kirk and Papachristos 2011). Neighborhood theorists, while noting the associations between disadvantage, immigration, demographic composition, and crime, have largely neglected
institutional and other social forces that may pattern both neighborhood ecology and subsequent crime and violence. For example, Clear et. al. (2003) state that “theorists of social disorganization have not regarded the effects of public policies as important considerations for their models of public safety. Public policies were generally thought of as responses to crime, not antecedents of it… (35).” While residential segregation (Massey and Denton 1993) and other forms of systemic racism (Peterson and Krivo 2010; Sampson 2012) have been established as antecedents of community disadvantage, scholars within this tradition have fallen short in further specifying the impacts on crime of a broad array of social institutions that impact neighborhood structure and disadvantage. Focal among these in the realm of crime and justice is the institution of punishment, as punishment has been identified by social scientists as a key cog in the (re)production of American stratification (Western 2006, Wacquant 2009, Uggen and Wakefield 2010). This research suggests that punishment, particularly in a time marked by mass incarceration (Garland 2001), mass probation (Phelps 2017) and the recent proliferation of monetary sanctions such as fines, fees, and restitution (Harris 2016, Harris et. al. 2017), may play an antecedent role to neighborhood structure. In particular, how might community loads of punishment, in its spatially concentrated character (Simes 2018) and varied forms of incarceration, probation, and legal financial obligations (LFOs), impact concentrated disadvantage (Sampson and Groves 1987, Samson et. al. 1997)? Punishment could play a role in both a) directly patterning neighborhood crime rates and violence and b) indirectly impact crime and violence via the moderating influence of neighborhoods. Concentrated punishment could serve as an additional exogenous force within this theoretical system, akin to the theorized factors of
concentrated disadvantage and residential stability, that inhibits the ability of a
neighborhood to create informal social control (e.g., Clear et al. 2003), thereby leading to
higher rates of crime and violence.

This dissertation chapter places punishment within models of neighborhood effects
and crime, and examines how punishment operates within these theories of neighborhood
ecology, concentrated disadvantage, and crime. The chapter starts by placing punishment
in the context of scholarship on neighborhood effect theories of crime, and discusses
previous work that is suggestive of a link between punishment and concentrated
disadvantage. Then, I describe the data construction process, which describes the
construction of a unique panel dataset constructed for estimating the impact of
community-level punishment loads on both violence and concentrated disadvantage.
Then, through the workhorse of fixed effects panel models, I find support for both
incarceration and probation being disadvantage-inducing aspects within neighborhoods,
which in turn are tied to elevated rates of violence. Further, I show that this iatrogenic
pathway of punishment reduces the net incapacitation and deterrence impacts of
incarceration, while also augmenting the criminogenic direct impacts of probation. I close
by contextualizing these findings within the broader theoretical and empirical context, as
well as discuss the limitations of the current study and directions for future research.

**Literature Review**

*Models of Neighborhood Ecology and Crime*

Neighborhood ecology and crime theories explicate how informal social controls
are generated, patterned, and impact the ecological space in which individuals are
embedded, as well as how neighborhood characteristics, such as racial heterogeneity,
concentrated disadvantage, and immigrant concentration, pattern crime rates via the mediating force of social cohesion. Neighborhood ecology theories have their roots in the ecological tradition of the Chicago school (Park and Burgess 1925). Even earlier, DuBois (1899) placed emphasis on “the environment in which a Negro finds himself (25)” in his exploration of Black crime rates, emphasizing that social ecology “must have an immense effect on his thought and life, his work and crime, his wealth and pauperism (25).” Social (dis)organization theory departs from Park and Burgess’ (1925) concentric zone theory, which charted urban change in circular zones emanating from the city center. One of Shaw and McKay’s (1931) key insights was their emphasis of high variability of delinquency rates within racial groups and nationalities, and that “within the same type of social area, the foreign born and the natives, recent immigrant nationalities, and older immigrants produce very similar rates of delinquents (160, emphasis added).” This empirical observation brought the causal focus away from individual-level explanations for neighborhood crime rates, to explanations centered around the ecological characteristics of place (Kubrin and Weitzer 2001). In sum, this early work highlights two key facets of that inform modern neighborhood effects research: 1) the correlation of certain neighborhood characteristics (e.g., concentrated disadvantage) with crime rates, and that 2) these relationships are largely stable in the face of changes in demographic changes, emphasizing the independent causal importance of place-based characteristics in the formation of crime.

Neighborhood ecology theories maintain that socially organized communities are characterized by strong informal social control, stable social networks that facilitate and reinforce the existence of informal social control, and a consensus on how “to solve
common problems and reach common goals (Bursik 1988:521).” The efficacy of informal social control is the ability of a community to supervise and monitor criminal behavior, recognize strangers and neighbors, create movement-governing rules, and engage in guardianship behavior on behalf of the greater community. It is by disrupting or impeding these potential social relations by which neighborhood ecology and characteristics induce variation in community crime and violence. Various neighborhood characteristics have been empirically examined as a part of this body of research, including concentrated disadvantage, ethnic heterogeneity, residential stability, among others (e.g., Sampson 1987). Early empirical examinations of the original social (dis)organization model show support for its tenets (e.g., Sampson and Groves 1987; Lowenkamp et. al. 2003), and the model has since been expanded upon in terms of the conceptualization and measurement of the mediating force of informal social control. Sampson et al. (1997) note that although social cohesion and solidarity is necessary for informal social control, what is also necessary is the inclination to take action on behalf of one’s neighbors that is of key importance to their conceptualization of collective efficacy: “it is the linkage of mutual trust and the willingness to intervene for the common good that defines the neighborhood context of collective efficacy (919).” Using data from Chicago, they find that neighborhood-level factors of concentrated disadvantage, immigrant concentration, and residential (in)stability account for about 70% of the variance in collective efficacy. In turn, collective efficacy significantly reduces neighborhood violent crime, net of prior neighborhood homicide levels, and mediates a large portion of the association between the neighborhood-level measures and individual-level victimization, perceived violence, and neighborhood-level homicide.
Further scholarship has supported this collective efficacy model of neighborhood effects (e.g., Morenoff et al. 2001), and has also expanded both the scope of the neighborhood characteristics that pattern crime rates such as community organizations (e.g., Sharkey et al. 2017), as well as the addition of both social networks (Bellair and Browning 2010) and legal cynicism (Kirk and Papachristos 2011) as important mediating mechanisms along the causal pathway between neighborhood structure and crime.

**Punishment, Crime, and Concentrated Disadvantage**

While a large body of criminological scholarship has examined the direct effects of punishment (primarily incarceration) on state and county rates of crime (e.g., Levitt 1996), the role of punishment concentrations within the theories of neighborhood ecology has received less scholarly attention. A key exception to this is Rose and Clear (1998), who conceptualize the spatial concentration of incarceration as another destabilizing facet of neighborhood structure that can inhibit communities’ ability to form informal social controls, thereby elevating community rates of crime. They hypothesize that high rates of incarceration can break up key social networks by the systematic removal of individuals from communities, thereby harming community levels of social capital and cohesion. Clear et al. (2003), empirically examining some of Rose and Clear’s (1998) hypotheses, find nonlinear effects of incarceration concentration, with initial decreases in crime at lower levels of incarceration, followed by increases as the incarceration rate increased (see also Renauer et al. 2006). Additional scholarship has shown that concentrations of former prisoners can directly disrupt collective efficacy (Drakulich et al. 2012, Kirk 2009). Although the majority of punishment-crime research examines the impacts of incarceration, recent scholarship from Minnesota shows how county-level probation
loads, but not incarceration or LFO amounts, are associated with higher levels of violent injuries as recorded by hospital admissions data (Santaularia, Larson, and Uggen 2021). Although this sociological inquiry has laid some theoretical foundation for incarceration’s role as an aspect itself of neighborhood structure and its ties to community violence, research to date has not fully examined the links between punishment, neighborhood structure, and crime, and established how punishment may not only be a part of neighborhood structure, but also how it may exacerbate the pernicious effects of other aspects of neighborhood structure (e.g., concentrated disadvantage). Further, punishment in the form of probation or monetary sanctions has not been discussed in the context of these neighborhood effects.

Research on the connection between punishment and inequality and disadvantage suggests that punishment plays a role in structuring American inequality and stratification (Western 2006, Wacquant 2009, Uggen and Wakefield 2010). This body of scholarship suggests that punishment, even in its varied forms, can exacerbate crime and violence by impacting neighborhood levels of inequality and disadvantage. Punishment’s collateral consequences can impact the familial, economic, and political efficacies of communities to challenge structures of disadvantage. Incarceration can lead directly to the breaking up of family structures (Comfort 2008, Apel 2016, Massoglia et. al. 2011), and parental incarceration has been tied to increased risk of child homelessness – particularly amongst African American children – and this relationship occurs, in part, via increases in family economic hardship as well as decreases in access to institutional support (Wildeman 2014). Punishment also has significant implications for material well-being, as punishment can beget increases in material insecurity and reliance on public assistance,
as well as decreases in earning potential (Schwartz-Soicher et al. 2011, Sugie 2012, Western 2002). Concentrations of punishment beget concentrations of criminal records (Shannon et al. 2017), which can disrupt access to the formal labor market (Pager 2003; Larson et al. 2021), restrict educational opportunities (Stewart and Uggen 2020; Bernburg et al. 2003), create housing instability (Harding et. al. 2013; Herbert et al. 2015), and dilute political power via felony disenfranchisement (Manza and Uggen 2003; Uggen, Larson, and Shannon 2016). The recent proliferation of LFOs (Harris et al. 2010) may also directly impact levels of disadvantage via predatory extraction of resources from already resource-deprived communities (Page and Soss 2021). At the community-level, tract-level LFO concentrations have been tied to future increases in poverty rates, marking monetary sanctions as a direct mechanism by which community disadvantage is maintained and perpetuated (O’Neill et al. 2021). Further, punishment, particularly probation, could weaken conventional ties to formal social institutions through “system avoidance.” For example, Brayne (2014) finds that those with criminal justice contact are less likely to interact with institutions such as hospitals, banks, employment, and schools, as compared to individuals without criminal justice contact. Thus, punishment may, in part, sever crucial ties to community institutions that affect both economic well-being and crime. Incarceration has also been tied to adverse health consequences, as well as rates of infant mortality (Schnittker and John 2007, Wildeman 2012), which could exacerbate disadvantage by spatially locating financial strain connected to health care costs alongside already existing forms of inequality. In turn, concentrated disadvantage has been one of the focal aspects of neighborhood structure that have been tied to neighborhood crime and violence rates (Sampson et. al 1997; Pratt and Cullen 2005;
Sampson 2012). Thus, concentrated disadvantage and resource deprivation may serve as a central mediating force between community-level punishment concentrations and crime and violence rates.

Research into the effects of punishment and crime has shown both crime reducing (e.g., Levitt 1996) and crime enhancing effects (e.g., Harding et. al. 2017) depending on the level of analysis and the research design and context. At the individual-level, incarceration has been shown to have crime-reducing benefits through incapacitation, but longer-term criminogenic effects (Harding et. al. 2017). At the aggregate level, previous research has found that incarceration has negative (Levitt 1996), positive (Clear et al 2003, Renauer et al. 2006) or null (Santaularia, Larson, and Uggen 2021) effects on violence. However, it is certainly plausible, if not probable, that in many contexts punishment may have both crime inducing and crime reducing capacities that net out to a summative effect of some direction, or could net to zero entirely. For example, research on the “incarceration ledger” finds that when comparing the lifesaving homicide returns to incarceration with the death-inducing effects of incarceration via infant mortality, the mortality returns of incarceration are weak, or even null (Light and Marshall 2018). Thus, it is possible that punishment could have crime-reducing direct effects, but this be offset, either fully or partially, by its countervailing aggravation of community disadvantage.

Uniting the above scholarship on neighborhood effects and the stratification effects of punishment, this second paper will explore the effects of community concentrations of incarceration, probation, and monetary sanctions on community-level rates of violence, while also examining the extent to which concentrated disadvantage mediates this relationship. Figure 7 depicts the theoretical model of punishment,
community disadvantage, and violence under investigation in the empirical work to follow.

Figure 7: Theoretical Model of Punishment, Disadvantage, and Violence

Data and Measures

This chapter leverages various forms of spatial and administrative data sources to build a unique panel data set to investigate the direct and indirect effects of punishment on crime and violence in Minneapolis, MN from 2011-2014. These data consist of spatial shapefiles from the Census Bureau’s Cartographic Boundary Shapefiles, Minnesota hospital discharge data from the Minnesota Hospital Association (MHA), court administrative data from the Minnesota State Court Administrator’s Office (SCAO), and sociodemographic data from the Census Bureau’s American Community Survey (ACS).

Spatial Data

In this second paper, the analysis of neighborhood effects requires a resolute analysis at the level of neighborhoods. However, the focal dependent variable utilized in this analysis, violent injuries as measured by the hospital discharge data, can only be
located at the level of a zip code. Therefore, to leverage the advantages of using the hospital discharge data as the key measure of crime and violence, this analysis will spatially situate these various sources of data at the zip code tabulation area (ZCTA)-year level in Minneapolis, Minnesota. I use the Census Bureau’s TIGER (U.S. Census Bureau 2022) ZCTA cartographic boundary shapefiles to define the spatial boundaries for each ZCTA in the United States, then spatially filter the ZCTA shapefiles using an intersection function – which defines the filter based on any shared common features (e.g., overlap, contiguity) based on Simple Feature (SF) geometries (Pevesma 2018) – with a shapefile of the city boundaries of Minneapolis (opendataminneapolis 2022).

I further spatially filter ZCTAs that have more than a 2% shared spatial intersection, to remove ZCTAs that primarily comprise neighboring cities and municipalities (e.g., St. Paul, MN). This process results in a base panel dataset of 88 observations, consisting of 22 Minneapolis ZCTAs over the years 2011-2014. The time range of the panel is indicative of the overlap present between the various forms of administrative data utilized in this analysis (see detailed data discussions below). All spatial shapefiles are accessed via the ‘tidyverse’ package, and all spatial transformations utilize functions from the `sf` package (Pevesma 2018), both of which are located within the R statistical computing environment (Walker 2022; Walker and Herman 2022). Finally, all spatial data are defined geographically using the North American Datum of 1983 (NAD83) coordinate reference system.

**Minnesota Hospital Discharge Data**

This study uses hospital administrative data hosted by the Minnesota Hospital Administration from 2011-2014 to measure community rates of violent injury. This
database includes both inpatient and outpatient data, and includes all patient encounters with health care providers and includes information on each record diagnosis and other characteristic characterized by the International Classification of Diseases codes (ICD-9). These data comprise the focal dependent variable of interest: the rate of violent injuries as recorded by hospital administrative data. The use of these data, over traditional sources of crime data, such as police records, is advantageous because 1) they are less amenable to recall biases to the extent that self-report victimization or police data are, and 2) avoids biases related to official tracking, enforcement, and reporting (Santaularia, Larson, and Uggen 2021). However, hospital records do systematically miss lower-levels of violence that do not rise to the level of an interaction with a health care provider, and do not capture non-violent crime events. With these caveats in mind, I argue that these data represent a more comprehensive picture of localized violent crime than other potential sources, and represent events that have serious consequences for community and individual well-being (Sharkey et al. 2012; Sharkey et al. 2014; Sharkey 2018a).

To create an indicator of violence at the ZCTA-year level, the MHA data was first filtered by ICD-9 code, and corresponding E- and V-codes. The ICD-9 codes represent the type of injury or condition that precipitated the health care provider visit, and the E- and V-codes modify that initial code for information on when/where on the body the injury occurred, as well as the history of the injury. Codes for violence were established following Santaularia et al. (2021), and include incidents of assault/homicide, rape, abuse, psychological abuse, among other forms of violence. Therefore, incidents of violence are delineated by the mechanism by which the injury occurred, differentiating injuries resulting from violence as compared to those via accident or another mechanism.
Each record was spatially located via each patient’s recorded residence zip code. The data was then aggregated by ZCTA and year, and were then spatiotemporally filtered to match the ZCTA-years in the base panel (see Spatial Data above). The focal dependent variable, the violent incidence rate per 1,000 residents, is expressed as follows:

\[ VR_{ij} = \left( \frac{VI_{ij}}{Pop_{ij}} \right) \times 1000 \]

Where \( VR_{ij} \) symbolizes the violence rate per 1,000 residents in a given ZCTA-year, \( VI_{ij} \) represents the count of violence events as recorded in each ZCTA-year by the MHA data, and \( Pop_{ij} \) is the ACS 5-year population estimate in the given ZCTA-year (see Sociodemographic Data below).

**Minnesota Court Administrative Data**

The key treatment variables in this study are measures of punishment from court administrative data received from the Minnesota Court Administrator’s Office. The SCAO data contains information on all criminal court events from 2004-2015. The spatial location of each case was determined by the defending party’s zip code, which spatially situates punishment where individuals live and carry out their sentences (apart from those incarcerated) as opposed to where the alleged offense occurred. This represents a significant advantage, as it spatially situated punishment where it is experienced. Specifically, I construct ZCTA-year level sums of sentenced incarceration days, sentenced probation days, and legal financial obligations (LFOs) - sentenced fines, fees, and restitution amounts by summing over case-level punishment amount measures. LFO amounts are converted to January 2018 USD to adjust for inflation over the panel.
All three focal punishment measures - incarceration, probation, and LFOs, are expressed per ZCTA-year resident, according to the following formulas:

\[
I_{ij} = \left( \frac{ID_{ij}}{Pop_{ij}} \right)
\]

\[
P_{ij} = \left( \frac{PD_{ij}}{Pop_{ij}} \right)
\]

\[
LFO_{ij} = \left( \frac{LFO_{ij}}{Pop_{ij}} \right)
\]

Where \(I_{ij}, P_{ij}, LFO_{ij}\) each represent per capita measures of incarceration, probation, and LFO punishment loads in each ZCTA-year respectively, \(ID_{ij}, PD_{ij}, LFO_{ij}\) represent the raw aggregated punishment incarceration days, probation days, and LFO dollars in each ZCTA-year respectively, and \(Pop_{ij}\) is the ACS 5-year population estimate in the given ZCTA-year (see Sociodemographic Data below).

**Sociodemographic Data**

I also merge onto the base panel various sociodemographic measures from the 5-year American Community Survey, which comprise both the key indicators for the construction of the latent construct of concentrated disadvantage (see Analytical Strategy below). Each ZCTA-year’s 5-year ACS estimates are accessed via the Census Bureau’s API via the `tidycensus` package in the R statistical computing environment (Walker and Herman 2022), and each year’s estimates represent “rolling window” estimates comprised of the focal years data combined with the previous four years, resulting in estimates for all ZCTAs in Minneapolis and more accurate ZCTA-year level sociodemographic characteristics as compared to the 1-year ACS estimates (Census
Bureau 2022). Each sociodemographic measure is expressed as a percentage of the total population. Similar to previous research on neighborhood effects (Sampson et al. 1997, Wodtke et al. 2011), I create a construct of concentrated disadvantage using time-varying indicators from the ACS using five indicators: the unemployment rate, percent below the poverty line, the percent of female headed-households, the percentage of the population with no high school diploma, and the percentage of Black residents. I also merge on various time-varying controls, to further isolate the impacts of the focal punishment measures. These include other measures of neighborhood structure that are popular in the neighborhood effects and crime research literature, such as percent foreign born, percent Hispanic, the residential mobility rate, the household ownership rate, and population density (see Jones and Pridemore 2018), as well as a broader set of measures of racial and age composition. Descriptive statistics for all variables used in the analysis can be found in Table 8.
Table 8: Descriptive Statistics for Panel

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHA Violence Rate</td>
<td>88</td>
<td>15.488</td>
<td>21.430</td>
<td>0.000</td>
<td>9.174</td>
<td>117.333</td>
</tr>
<tr>
<td>Incarceration Days per Capita</td>
<td>88</td>
<td>5.465</td>
<td>7.566</td>
<td>0.121</td>
<td>3.001</td>
<td>48.494</td>
</tr>
<tr>
<td>Probation Days per Capita</td>
<td>88</td>
<td>10.887</td>
<td>20.418</td>
<td>2.252</td>
<td>13.744</td>
<td>127.488</td>
</tr>
<tr>
<td>LFO per Capita</td>
<td>88</td>
<td>45.681</td>
<td>23.912</td>
<td>7.352</td>
<td>49.929</td>
<td>136.630</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>88</td>
<td>10.621</td>
<td>6.942</td>
<td>2.252</td>
<td>7.938</td>
<td>36.492</td>
</tr>
<tr>
<td>Poverty Rate</td>
<td>88</td>
<td>19.905</td>
<td>12.261</td>
<td>0.000</td>
<td>18.912</td>
<td>43.527</td>
</tr>
<tr>
<td>Percent Female Headed Household</td>
<td>88</td>
<td>4.401</td>
<td>2.589</td>
<td>0.000</td>
<td>4.257</td>
<td>11.721</td>
</tr>
<tr>
<td>Percent No HS Diploma</td>
<td>88</td>
<td>7.152</td>
<td>4.780</td>
<td>0.000</td>
<td>6.168</td>
<td>19.828</td>
</tr>
<tr>
<td>Percent Black</td>
<td>88</td>
<td>16.578</td>
<td>14.243</td>
<td>0.000</td>
<td>11.340</td>
<td>53.843</td>
</tr>
<tr>
<td>Percent Foreign</td>
<td>88</td>
<td>15.468</td>
<td>8.971</td>
<td>5.598</td>
<td>12.871</td>
<td>45.417</td>
</tr>
<tr>
<td>Residential Mobility</td>
<td>88</td>
<td>20.745</td>
<td>16.457</td>
<td>12.000</td>
<td>25.295</td>
<td>86.737</td>
</tr>
<tr>
<td>Percent Own III</td>
<td>88</td>
<td>19.556</td>
<td>19.650</td>
<td>0.000</td>
<td>20.916</td>
<td>36.555</td>
</tr>
<tr>
<td>Population Density (Z)</td>
<td>88</td>
<td>0.000</td>
<td>1.000</td>
<td>-1.571</td>
<td>-0.173</td>
<td>3.098</td>
</tr>
<tr>
<td>Percent Hispanic</td>
<td>88</td>
<td>7.221</td>
<td>5.584</td>
<td>0.785</td>
<td>6.228</td>
<td>25.562</td>
</tr>
<tr>
<td>Percent Native American</td>
<td>88</td>
<td>1.163</td>
<td>1.179</td>
<td>0.000</td>
<td>0.829</td>
<td>4.818</td>
</tr>
<tr>
<td>Percent Non-Hispanic White</td>
<td>88</td>
<td>63.977</td>
<td>19.637</td>
<td>15.054</td>
<td>68.762</td>
<td>90.708</td>
</tr>
<tr>
<td>Percent Asian</td>
<td>88</td>
<td>7.569</td>
<td>7.913</td>
<td>1.428</td>
<td>4.061</td>
<td>38.695</td>
</tr>
<tr>
<td>Percent Hawaiian/Pac. Isl.</td>
<td>88</td>
<td>0.057</td>
<td>0.110</td>
<td>0.000</td>
<td>0.000</td>
<td>0.660</td>
</tr>
<tr>
<td>Percent Other Race</td>
<td>88</td>
<td>0.566</td>
<td>1.559</td>
<td>0.000</td>
<td>0.672</td>
<td>8.800</td>
</tr>
<tr>
<td>Percent Biracial</td>
<td>88</td>
<td>7.003</td>
<td>2.406</td>
<td>2.592</td>
<td>6.616</td>
<td>12.886</td>
</tr>
<tr>
<td>Percent Age Below 18</td>
<td>88</td>
<td>17.455</td>
<td>9.090</td>
<td>0.000</td>
<td>18.488</td>
<td>37.000</td>
</tr>
<tr>
<td>Percent Age 19-20</td>
<td>88</td>
<td>30.257</td>
<td>20.336</td>
<td>10.086</td>
<td>22.767</td>
<td>99.051</td>
</tr>
<tr>
<td>Percent Age 30-40</td>
<td>88</td>
<td>28.450</td>
<td>8.793</td>
<td>0.000</td>
<td>31.080</td>
<td>44.988</td>
</tr>
<tr>
<td>Percent Age 50-60</td>
<td>88</td>
<td>18.193</td>
<td>5.975</td>
<td>0.000</td>
<td>19.250</td>
<td>26.691</td>
</tr>
<tr>
<td>Percent Age 70+</td>
<td>88</td>
<td>5.644</td>
<td>2.683</td>
<td>0.000</td>
<td>5.931</td>
<td>10.700</td>
</tr>
</tbody>
</table>

Analytical Strategy

Spatial Mapping and Bivariate Correlations

The first step of the analysis is the construction of choropleth maps of the key outcome and treatment variables: the violence rate, incarceration per capita, probation per capita, and LFOs per capita. These maps will display the spatial variation in ZCTA-level violence rates across Minneapolis. I then calculate simple Pearson product-moment correlations between each measure, to assess the first-order associations between these punishment measures and violence.

Confirmatory Factor Analysis

The second aspect of the analysis is the creation of the latent neighborhood construct of concentrated disadvantage via confirmatory factor analysis (CFA)
measurement models. CFA is a theory-driven technique where proposed relationships between observed indicator variables, and an unobserved latent variable that is purported to account for the dependencies between the observed variables, is specified. The coefficients, or factor loadings, of these ACS items are estimated via maximum-likelihood estimation (these are also known as reliability coefficients in their squared form). The model scales the dependent variable, concentrated disadvantage, by assuming that the latent variable’s variance is standardized to 1, thereby making the factor loadings represent standardized effects between the latent construct and a particular indicator variable. I then leverage predictions from the measurement model to create measures of concentrated disadvantage used in the subsequent panel modeling.

Many social science studies using multiple measures as indicators of a latent construct will create a simple summative measure of the indicators, and then use the combined measure in subsequent analysis. This has the limitation of treating the measurement error involved in the creation of the measure as variation in the measure itself, and the use of CFA avoids this limitation by explicitly accounting for measurement error within the measurement model. Modification indices, which measure the improvement in model $\chi^2$ goodness-of-fit for additional model specifications, were used to specify shared covariances between the observed indicator variables.

CFA model fit is captured through four standard measures of model fit: 1) A likelihood-ratio chi-squared test between the fitted model and a saturated model, which is the model that fits the covariances perfectly, 2) the Root Mean Square Error of Approximation, a scaled absolute measure of fit that adjusts the chi-squared for model parsimony (values below .05 indicate good model fit; PCLOSE is also reported, the
probability that RMSEA is below .05), 3) the Comparative Fit Index (CFI), which compares the fit of the null model and the fitted model adjusting for model complexity (values > .9 indicate good model fit) and 4) the Standardized Root Mean Square Residual (SRMR) which quantifies the standardized difference between the observed covariances and the predicted covariances (values < .08 indicate good model fit) (Hu and Bentler 1998). The CFA model for concentrated disadvantage was fit using the ‘lavaan’ package in R (Rosseel 2012).

Two-Way Fixed-Effects Panel Model

To evaluate the impact of punishment on both neighborhood structure and subsequent crime, I estimate two-way fixed effect panel models. The panel structure of the data allows for the control of unobserved heterogeneity across both neighborhoods and time, which aids in alleviating endogeneity concerns due to omitted-variable bias. Effectively, these models control for time-constant heterogeneity between ZCTAs, as well as over time heterogeneity common to all ZCTAs. The first specifications will test whether the forms of punishment impact ZCTA-level concentrated disadvantage, the proposed first-leg of the relationships depicted in Figure 1.

\[
CD_{ij} = \beta_{1}I_{ij,t-1} + \lambda X_{ij} + \theta_i + T_j + \epsilon_{ij}
\]

\[
CD_{ij} = \beta_{2}P_{ij,t-1} + \lambda X_{ij} + \theta_i + T_j + \epsilon_{ij}
\]

\[
CD_{ij} = \beta_{3}LFO_{ij,t-1} + \lambda X_{ij} + \theta_i + T_j + \epsilon_{ij}
\]

where \(CD_{ij}\) represents the standardized measure of concentrated disadvantage constructed from the measurement model (see Table 9), \(\lambda X_{ij}\) represents observed, time-varying control variables of neighborhood structure and composition from the ACS, and
\( \varepsilon_{ij} \) represents unobserved time-varying influences on the focal aspect of neighborhood structure – concentrated disadvantage. The focal coefficients, \( \beta_1, \beta_2, \beta_3 \), represent the impacts of each type of punishment concentration, net of the time-varying controls and fixed effects, on within-ZCTA changes in concentrated disadvantage. I lag each of these punishment measures due to the endogenous nature of the simultaneous relationship between punishment and concentrated disadvantage, as previous research has linked punishment levels to concentrated disadvantage and inequality (Woolredge 2007, Wacquant 2009). The ZCTA fixed effects (\( \theta_i \)) capture average, time-stable, unobserved influences on each ZCTA’s structure (and effectively control for time-constant preexisting ZCTA structure) and year fixed effects (\( T_j \)) capture average unobserved influences common across neighborhoods within each year, such as city-wide policy changes in Minneapolis.

To examine the direct effects of both punishment and concentrated disadvantage, as well as examine the extent to which concentrated disadvantage mediates the effects of punishment, I model the violence rate per 1,000 residents as measured by the MHA hospital data. Due to the highly correlated nature of the forms of punishment (see Figure 8), I create separate models for each form of punishment per capita, and with each focal punishment measure estimate two specifications. The first specifications can be notated as,

\[
V_{ij} = \beta_1 I_{ij,t-1} + \lambda X_{ij} + \theta_i + T_j + \varepsilon_{ij}
\]

\[
V_{ij} = \beta_2 P_{ij,t-1} + \lambda X_{ij} + \theta_i + T_j + \varepsilon_{ij}
\]

\[
V_{ij} = \beta_3 LFO_{ij,t-1} + \lambda X_{ij} + \theta_i + T_j + \varepsilon_{ij}
\]
where $V_{ij}$ represents the violence rate per 1,000 as measured by the MHA hospital data, $\lambda X_{ij}$ represent observed time-varying covariates from the ACS, and $\varepsilon_{ij}$ represent unobserved time-varying influences on the neighborhood’s crime rate. The focal coefficients, $\beta_1, \beta_2, \beta_3$, represent the impacts of each type of punishment concentration on the rate of violent injury. I also lag each of these measures due to the endogenous nature of the simultaneous relationship between punishment and violence (e.g., Levitt 1996).

The ZCTA fixed effects ($\theta_i$) capture average, time-stable, unobserved influences on each ZCTA’s violence rate (and effectively control for preexisting ZCTA levels of violence) and year fixed effects ($T_j$) capture average unobserved influences common across neighborhoods within each year, such as city-wide policy changes. The full violence rate specifications are as follows:

$$V_{ij} = \beta_1 I_{ij,t-1} + \pi CD_{ij} + \lambda X_{ij} + \theta_i + T_j + \varepsilon_{ij}$$

$$V_{ij} = \beta_2 P_{ij,t-1} + \pi CD_{ij} + \lambda X_{ij} + \theta_i + T_j + \varepsilon_{ij}$$

$$V_{ij} = \beta_3 M_{ij,t-1} + \pi CD_{ij} + \lambda X_{ij} + \theta_i + T_j + \varepsilon_{ij}$$

which adds the term, $\pi CD_{ij}$, which represents the effects of concentrated disadvantage. These specifications will allow me to a) determine if any direct effects of punishment remain after accounting for neighborhood structure, and b) how much of the impact of punishment on subsequent crime rates is mediated by concentrated disadvantage. All fixed-effect specifications are estimated with demeaned data via the within transformation, which removes both time-stable neighborhood variation and neighborhood-stable time variation from data on both sides of the equation, leaving the estimation of parameters to rely upon time-varying variation within ZCTAs (see
Woolridge, 2016). Because my analytic sample includes the entire population of all ZCTA-years during the study timeframe, interpretive emphasis will be placed on the magnitude of the estimated effects, as opposed to measures of statistical inference (i.e., p-values). All fixed effects specifications are estimated using the ‘plm’ package in R (Croissant and Millo 2008).

**Mediation Analysis**

Using these specifications in concert with the estimates from the neighborhood structure models, I then perform a mediation analysis for each punishment variable, to estimate both the direct and indirect effects, via concentrated disadvantage, of punishment using a bootstrapping method that is useful for smaller sample sizes (Imai et al. 2010). Essentially, this method computes estimates of the direct and indirect effects from the estimated fixed effects specifications over 1,000 bootstrapped samples, resulting in point estimates of both the direct and indirect effects, as well as variance estimates for inference.\(^{20}\) This method estimates the average causal mediation effect (ACME) - the indirect effect of each form of punishment on violence that goes through concentrated disadvantage, the average direct effect – the effect of punishment net of concentrated disadvantage, and the total effect - the sum of the direct and indirect effects. All mediation analyses are estimated via the ‘mediation’ package in R (Tingley et al. 2014).

**Results**

The first component of this chapter’s analysis examines the spatial distribution, as well as the bivariate associations between the focal punishment measures and the

---

\(^{20}\) This analysis uses population data, and therefore the point estimates are of primary interest in the mediation analysis.
violence rate. Starting in the upper left panel of Figure 8, the violence rate amongst residents is not distributed across space equally, with higher rates in the North Minneapolis and Phillips neighborhoods, alongside a downtown ZCTA, marked in red, that stands out as an extreme observation. In comparison the ZCTAs neighboring the southwestern suburbs, as well as the eastern border with St. Paul, MN exhibit markedly smaller rates of violent injury amongst residents. This pattern is consistent, albeit on separate scales, with incarceration capita and probation capita, as evidenced in the upper right and lower left panes of Figure 8 respectively. In contrast, LFO capita, in the lower right pane of Figure 8, shows a slightly different spatial pattern, with LFOs per capita spread a bit more evenly across the ZCTAs in Minneapolis. However, the ZCTAs with the highest LFOS per resident do mirror the patterns shown in the other three plots. These patterns are suggestive of robust correlations of punishment and violence.

\[\text{ZCTA 55402 is consistently higher across the choropleths, which is driven by a small denominator, as ZCTA 55402 has significantly lower total population on average. ZCTA 55402 has a mean total population across 2011-2014 of 402.75 as compared to the other ZCTAs with a mean of 20,730.45.}\]

\[\text{It should be noted that these simple correlations suffer from simultaneity bias, as the link between punishment and violence contains both the responses of violence in response to the effects of punishment, as well as punishment’s responses to levels of violence.}\]
Consistent with the maps, simple bivariate correlations reveal significant positive associations of the violence rate with all three forms of punishment per capita.

Specifically, incarceration capita and the violence rate exhibit a bivariate association of .68, slightly larger than the violence rate’s associations with probation capita (.64), and LFO capita (.66), albeit all correlations are very similar. This provides preliminary evidence that punishment and violence tend to positively covary in Minneapolis.

Commensurate with these associations, the bivariate associations between the punishment measures show markedly high correlations. The bivariate association between incarceration capita and probation capita is a high .95, and incarceration’s correlation with LFO capita is .91. Similarly, the bivariate association between probation capita and
LFO capita is large at .93. This indicates that different forms of punishment are highly related to each other in space, and places with large amounts of per capita incarceration also have large amounts of per capita probation and monetary sanctions. Overall, these simple analyses suggest two key findings moving forward: that levels of violence do covary, at least at the bivariate level, with punishment loads, and that all three legs of punishment are highly correlated across the panel spatiotemporally.

Table 9 presents the CFA measurement of concentrated disadvantage, using five indicators from the 5-year ACS estimates at the ZCTA-year level: the unemployment rate, percent below the poverty line, the percent of female headed-households, the percentage of the population with no high school diploma, and the percentage Black. An initial model specifying only factor loadings (FL) between the latent measure of concentrated disadvantage and the indicators was estimated, but modification indices—which estimate the improvement in model $\chi^2$ as a result of further model specifications—indicated that a shared covariance specification between percent Black and the unemployment rate would significantly improve model fit. The final model adds in this covariance specification, which suggests that the unexplained portions of both percent Black and the unemployment rate shared a common source.

<table>
<thead>
<tr>
<th>LHS</th>
<th>Specification</th>
<th>RHS</th>
<th>Std(Beta)</th>
<th>SE</th>
<th>P.Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coac. Dis.</td>
<td>FL</td>
<td>Unemp. Rate</td>
<td>0.240</td>
<td>0.114</td>
<td>0.035</td>
</tr>
<tr>
<td>Coac. Dis.</td>
<td>FL</td>
<td>Poverty Rate</td>
<td>0.598</td>
<td>0.066</td>
<td>0</td>
</tr>
<tr>
<td>Coac. Dis.</td>
<td>FL</td>
<td>Female-HH Rate</td>
<td>0.745</td>
<td>0.049</td>
<td>0</td>
</tr>
<tr>
<td>Coac. Dis.</td>
<td>FL</td>
<td>No HS Diploma Rate</td>
<td>0.811</td>
<td>0.039</td>
<td>0</td>
</tr>
<tr>
<td>Coac. Dis.</td>
<td>FL</td>
<td>Percent Black</td>
<td>1.008</td>
<td>0.025</td>
<td>0</td>
</tr>
<tr>
<td>Unemp. Rate</td>
<td>Cor.</td>
<td>Percent Black</td>
<td>1.823</td>
<td>3.077</td>
<td>0.554</td>
</tr>
</tbody>
</table>

$LR\chi^2$ vs. saturated (4) = 1.909, RMSEA = .00 (PCLOSE = .819), CFI = 1.0, SRMR = .024
The CFA measurement model, using five different indicators of the latent variable of concentrated disadvantage, all exhibit statistically significant factor loadings onto the latent variable, and the correlations are consistent with research documenting the institutional forces (e.g., redlining, racial covenants) undergirding structural racism that has concentrated race and disadvantage in both Minneapolis and across the United States (e.g., Mapping Prejudice Project 2022; Massey and Denton 1993). These measures represent the standardized effect of the latent variable on each of the indicator variables, with percent Black (1.0) and the no HS diploma rate (.81) having the strongest loadings onto concentrated disadvantage, followed by the female headed household rate (.75), poverty rate (.60), and unemployment rate (.24) respectively. Although the covariance term between percent Black and the unemployment rate is not statistically significant, the estimated effect is high (1.82), and its inclusion in the model significantly improves the model fit.

Overall, all measures of model fit are indicative of excellent model fit. The statistically nonsignificant LR $\chi^2$ test is indicative of no significant differences between the specified model and a perfectly fitting saturated model. In addition, both the absolute fit measure of RMSEA (.00, probability < .05 .82), and the comparative CFI (1.0), both of which penalize for model complexity, are both indicative of superb model fit. Finally, an SRMR of .024, lower than .08, is indicative of very few quantifiable standardized discrepancies between the observed and predicted correlation matrices. Moving forward, linear predictions from this measurement model will serve as the standardized measure of concentrated disadvantage used in the fixed effect panel models below.
Table 10 displays fixed effect panel models regressing concentrated disadvantage on each form of punishment load, alongside the time-varying sociodemographic controls. Model 1 indicates a positive association between lagged incarceration capita and concentrated disadvantage. Specifically, a one day increase in lagged incarceration per resident is associated with a .016 standard deviation increase in concentrated disadvantage net of other factors, a fairly large effect in magnitude for a day per capita increase. Equivalently, a one standard deviation increase (7.6) in lagged incarceration per capita is associated with a .12 (7.6*.016) standard deviation increase in concentrated disadvantage.

Similarly in Model 2, lagged probation per capita also exhibits a positive association with concentrated disadvantage, with a one-day per capita increase in lagged probation associated with a .004 standard deviation increase in concentrated disadvantage. This raw effect is smaller than that of the lagged incarceration per capita effect, but the explanatory variable of lagged probation per capita exhibits greater variation (see Table 8). In terms of a standardized effect, a one standard deviation increase in lagged probation per capita (20.4) is associated with a .08 (20.4*.004) standard deviation increase in concentrated disadvantage. The incarceration effect is modestly larger than the probation effect, which could be due to the removal of potential income earners from neighborhoods (e.g., Clear 2003).

In contrast to the positive impacts of both lagged incarceration capita and probation capita, Model 3 shows a negative impact of lagged LFOs per capita on concentrated disadvantage, with a one-dollar per capita increase in LFOs associated with a .004 standard deviation decrease in concentrated disadvantage. In terms of a
standardized effect, as LFOs per capita exhibit greater variation than either of the other punishment per capita measures, a one standard deviation increase (23.9) in lagged LFOs per capita is associated with a -.09 decrease in concentrated disadvantage. This finding is in contrast to recent research that finds a positive association with LFO loads and future poverty rates (O’Neill et. al. 2021), which comprise one aspect of the broader construct of concentrated disadvantage. However, evidence from courtroom ethnographic observations in Hennepin County (the county in which Minneapolis is located) suggest that judges are adept at taking into account abilities to pay, even waiving mandatory state fees and surcharges that judges in other counties do not (Harris et a. 2017), which could offer an explanation why LFOs, often burdensome on those under their purview (Harris et al. 2010), do not play as large a role in increasing disadvantage in Minnesota. I further contextualize and discuss this finding in the Discussion and Conclusion section below.
Table 10: FE Panel Models of Concentrated Disadvantage, 2011-2014

<table>
<thead>
<tr>
<th>Concentrated Disadvantage</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incar. Capita (t-1)</td>
<td>0.016</td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>Prob. Capita (t-1)</td>
<td>0.004</td>
<td>(0.004)</td>
<td></td>
</tr>
<tr>
<td>LFO Capita (t-1)</td>
<td>-0.004</td>
<td>(0.004)</td>
<td></td>
</tr>
<tr>
<td>Percent Hispanic</td>
<td>-0.086***</td>
<td>(0.017)</td>
<td>-0.088***</td>
</tr>
<tr>
<td>Percent Native American</td>
<td>-0.157***</td>
<td>(0.041)</td>
<td>-0.168***</td>
</tr>
<tr>
<td>Percent White</td>
<td>-0.116***</td>
<td>(0.018)</td>
<td>-0.117***</td>
</tr>
<tr>
<td>Percent Asian</td>
<td>-0.119***</td>
<td>(0.027)</td>
<td>-0.102**</td>
</tr>
<tr>
<td>Percent HPI</td>
<td>-0.195</td>
<td>(0.170)</td>
<td>-0.258</td>
</tr>
<tr>
<td>Percent Other Race</td>
<td>-0.172</td>
<td>(0.129)</td>
<td>-0.187</td>
</tr>
<tr>
<td>Percent Biracial</td>
<td>-0.019</td>
<td>(0.015)</td>
<td>-0.017</td>
</tr>
<tr>
<td>Percent Foreign Born</td>
<td>-0.015</td>
<td>(0.019)</td>
<td>-0.019</td>
</tr>
<tr>
<td>Residential Mobility Rate</td>
<td>0.003</td>
<td>(0.011)</td>
<td>0.013</td>
</tr>
<tr>
<td>Own HH Rate</td>
<td>0.011</td>
<td>(0.028)</td>
<td>0.008</td>
</tr>
<tr>
<td>Population Density (Z)</td>
<td>-0.239</td>
<td>(0.220)</td>
<td>-0.281</td>
</tr>
<tr>
<td>Percent 18 or Below</td>
<td>-0.066</td>
<td>(0.048)</td>
<td>-0.049</td>
</tr>
<tr>
<td>Percent 19-29</td>
<td>-0.022</td>
<td>(0.044)</td>
<td>-0.020</td>
</tr>
<tr>
<td>Percent 30-49</td>
<td>-0.033</td>
<td>(0.041)</td>
<td>-0.055</td>
</tr>
<tr>
<td>Percent 50-69</td>
<td>-0.053</td>
<td>(0.038)</td>
<td>-0.045</td>
</tr>
<tr>
<td>ZCTA FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>66</td>
<td>65</td>
<td>66</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.826</td>
<td>0.822</td>
<td>0.823</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.564</td>
<td>0.554</td>
<td>0.557</td>
</tr>
</tbody>
</table>

Notes: ***Significant at the 0.1 percent level. **Significant at the 1 percent level. *Significant at the 5 percent level.

Table 11 relays fixed effect panel models of the violence rate. Models 1-3 regress the violence rate on each of the lagged punishment measures and other controls, and models 4-6 add to the models the standardized measure of concentrated disadvantage, to
examine the potential mediating role of community disadvantage. I begin by discussing the effects of the focal lagged punishment measures in Model 1-3, then discuss how each effect is altered upon the introduction of concentrated disadvantage into the model, as well as the estimated effects of concentrated disadvantage on the ZCTA-level violence rate.

Model 1 shows a large negative effect of lagged incarceration capita on the violence rate, with a one unit increase of lagged incarceration associated with a .84 decrease in violent hospital discharges per 1,000 residents, net of controls and other aspects of neighborhood structure. This short-term effect of incarceration is consistent with previous research on the incapacitation function of incarceration (e.g., Harding et al. 2017). In contrast, Model 2, which regresses the violence rate on lagged probation per capita and other controls, is similar in magnitude (for a similar unit increase), but opposite of lagged incarceration capita directionally. Specifically, a one day increase in lagged probation capita is associated with a .86 increase in the neighborhood violence rate per 1,000 residents. This marks probation loads as a potential iatrogenic force in the perpetuation of neighborhood violence, and it is consistent with previous research in Minnesota at the county-level (Santaularia, Larson, and Uggen 2021; see also Worrall et al. 2004).

Finally, Model 3 indicates a negative association between lagged LFOs per capita and the violence rate, with a one dollar increase in lagged LFO loads corresponding to a .27 decrease in violent hospital discharges per 1,000 residents. For a standard deviation increase in any of these lagged punishment measures, LFOs are comparable in magnitude with roughly a 6.4 decrease in violent discharges per 1,000 residents (23.9*-.266 = -
6.36), as compared to incarceration loads (7.57*-837 = -6.3). In contrast, probation is not only directionally opposite these other punishment effects, but also stronger in magnitude, as a one standard deviation increase in probation is associated with a 17.5 (20.4*-859 = 17.5) increase in violent hospital discharges per 1,000 community residents.

In sum, Models 1-3 establish the overall, short-term, bifurcated effects of punishment loads on community violence, with incarceration and LFOs associated with less violence on average, and probation associated with higher levels of violence.

<table>
<thead>
<tr>
<th>Table 11: FE Panel Models of the Violence Rate, 2011-2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence rate per 1,000</td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>Incar. Capita (t-1)</td>
</tr>
<tr>
<td>Prob. Capita (t-1)</td>
</tr>
<tr>
<td>LFO Capita (t-1)</td>
</tr>
<tr>
<td>Concentrated Disadvantage</td>
</tr>
<tr>
<td>Percent Hispanic</td>
</tr>
<tr>
<td>Percent Native American</td>
</tr>
<tr>
<td>Percent White</td>
</tr>
<tr>
<td>Percent Asian</td>
</tr>
<tr>
<td>Percent HPI</td>
</tr>
<tr>
<td>Percent Other Race</td>
</tr>
<tr>
<td>Percent Bi racial</td>
</tr>
<tr>
<td>Percent Foreign Born</td>
</tr>
<tr>
<td>Residential Mobility Rate</td>
</tr>
<tr>
<td>Own HH Rate</td>
</tr>
<tr>
<td>Population Density (Z)</td>
</tr>
<tr>
<td>Percent 18 or Below</td>
</tr>
<tr>
<td>Percent 19-29</td>
</tr>
<tr>
<td>Percent 30-49</td>
</tr>
<tr>
<td>Percent 50-69</td>
</tr>
<tr>
<td>ZCTA FE</td>
</tr>
<tr>
<td>Year FE</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>R²</td>
</tr>
<tr>
<td>Adjusted R²</td>
</tr>
</tbody>
</table>

Notes:
*Significant at the 0.1 percent level.
**Significant at the 1 percent level.
*Significant at the 5 percent level.
Models 4-6 in Table 11 represent fixed effects panel specifications, similar to the first three models, but add in the standardized measure of concentrated disadvantage constructed from the measurement model in Table 9. First, it is clear that concentrated disadvantage has a strong, robust relationship with community violence: the effect of concentrated disadvantage is consistently large and positive, with a one standard deviation increase in disadvantage being associated with 37.9, 27.3, 32.1 increases in violent hospital discharges per 1,000 in each respective model. This is consistent with previous neighborhood effects research that marks concentrated disadvantage as a central force in patterning neighborhood crime and violence (Sampson et al. 1997; Sampson 2012).

These full specifications also allow the examination of the potential mediating role of concentrated disadvantage. In Model 4, the effect of incarceration capita upon introduction of concentrated disadvantage to the model becomes larger, with a one unit increase in lagged incarceration days per capita associated with a reduction of 1.4 violent injuries per 1,000. This is indicative of concentrated disadvantage acting as a “suppressor” variable, which occurs when the direct and mediated effects of a treatment variable differ in sign (Tzelgov and Henik 1991). Essentially, incarceration per capita has two pathways to impacting violence: a negative direct pathway, and a positive indirect one via concentrated disadvantage as evidenced by the positive coefficient in Table 10 Model 1 alongside the large impact of concentrated disadvantage in Table 11 Model 4. Effectively, Model 1’s lagged incarceration capita coefficient estimate “combines” the two bifurcating pathways, thereby showing a reduced effect when not conditioning on concentrated disadvantage. Then, once disadvantage is conditioned upon, the direct
pathway is extricated from the positive indirect pathway via concentrated disadvantage. This finding represents the double-edged sword of incarceration, as increases do appear to reduce violence in the short-term through incapacitation and other potential mechanisms (e.g., deterrence), but that these reductions are partially offset by increasing neighborhood disadvantage, and denotes an iatrogenic connection between violence and incarceration.

Models 5 and 6 provide evidence towards more classic mediation relationships, as the initial effects estimated in Models 2 and 3 attenuate partially towards zero upon the inclusion of concentrated disadvantage into the specification. In Model 5, the effect of lagged probation capita is reduced from .86 to .75, as probation also has a positive pathway to violence via concentrated disadvantage. This indicates that part of probation’s criminogenic effect is mediated via concentrated disadvantage, and that community violence rates may be lower if the disadvantage inducing properties of probation were altered. Similarly, the negative impact of LFOs on violence is attenuated (-.27 to -.13) as LFOs have a negative pathway to the violence rate via concentrated disadvantage. This finding, combined with the results of Model 3 in Table 10, suggest that LFOs play a small, but measurable role in the reduction of disadvantage, which in turn is associated with lower rates of violence. I provide a detailed contextualization and discussion of these varied effects of punishment loads on violence in the discussion and conclusion section.

To corroborate the findings from the models in Table 4, I also estimate both direct and indirect effects using a bootstrapped mediation analysis, which computes direct and indirect estimates using 1,000 bootstrapped samples from the panel data. Figure 9 plots
the mediation analysis of incarceration capita, with an average causal direct effect (ACME) via concentrated disadvantage of .61, accompanied by an average direct effect (ADE) of -1.43. Together, the total effect, which sums the ACME and ADE, of incarceration per capita is -.83. Thus, the short-term incapacitation effects (e.g., Harding et. al. 2017) of incarceration loads are combatted by incarceration capita’s ability to enhance disadvantage. This represents key evidence towards a *bifurcating* effect of incarceration in both a negative direct effect, partially offset by a positive indirect effect via concentrated disadvantage. Figure 10 plots the estimated direct and indirect effects of lagged probation capita, with an estimated ACME of .14, ADE of .61, and a total effect of .75. This mediation analysis indicates that about 18.5% of the criminogenic effect of neighborhood probation loads operate through the mediator of concentrated disadvantage, and that the remaining 81.5% of the effect is direct. This establishes probation, alongside incarceration, as an iatrogenic force of punishment on violence and crime through the (re)production of community disadvantage. Finally, Figure 11 plots the bootstrap estimated effects of lagged LFO amounts per capita, resulting in an ACME of -.002, ADE of -.13, and total effect of .13. The mediation analysis of LFOs indicates that the average effects of LFOs, both directly and indirectly through concentrated disadvantage, are smaller than the other forms of punishment, but are both negative in direction. Specifically, about a little over 1 percent of the negative effect of LFOs operates through reducing disadvantage, indicating that the effect of LFOs, albeit smaller, operates primarily via other mechanisms.
Discussion and Conclusion

This empirical dissertation chapter set out to investigate the impacts of punishment concentrations in its varied forms directly on community rates of violence, as well as indirectly via the pathway via the (re)structuring of neighborhoods – using concentrated disadvantage as a neighborhood characteristic responsive to punishment. Setting punishment as an exogenous force in theories of neighborhood ecology and crime, this chapter makes both a theoretical and empirical contribution to both the vast neighborhoods and crime literatures, as well as empirically extend the implications inherent in the collateral consequences and punishment and stratification bodies of scholarship. This research not only has important implications for future research on the nexus of punishment, disadvantage, and crime at the community-level, but also holds implications for community public safety policies.
First, this chapter confirms with data in Minneapolis, MN the heavy spatial concentration of incarceration (Simes 2018; Sampson and Loeffler 2010) into specific neighborhoods. It also documents how patterns in the concentrations of probation and monetary sanctions follow closely the spatiotemporal patterns in custodial sentencing. Second, this analysis revealed the positive relationships between incarceration and probation concentrations and concentrated disadvantage, as well as the unexpected small, but negative, relationship between LFO loads and community disadvantage. The effects of incarceration and probation align with tenets in the punishment and stratification literature (e.g., Uggen and Wakefield 2010), and are roughly similar in standardized magnitude, indicating that while incarceration is often assumed to carry the greatest consequences, probation is “holding its own” in terms of patterning neighborhood changes in disadvantage and inequality.

The negative relationship between LFO loads and concentrated disadvantage is a bit puzzling, and is inconsistent with research tying community level monetary sanction loads to future poverty levels (O’Neill et al. 2021). However, it is consistent with the spatial variation in Figure 8, as LFOs in Minnesota are more evenly distributed across areas with varying levels of disadvantage as compared to the incarceration and probation amounts. This is likely due to offenses such a DUIs, which carry large fines and fees, being more common in more affluent areas. Additionally, Minnesota has been shown to have relatively lower fine and fee amounts as compared to other states (Harris et al. 2017), and generally has less deleterious nonpayment consequences such as reincarceration. Further, courtroom ethnographic evidence suggests that Hennepin County, in which Minneapolis, MN resides, is particularly adept at taking into account
defendant ability to pay, even waiving mandatory state fees and surcharges for indigent defendants (Harris et al. 2017), which may have played a role in lessening the deleterious disadvantage impacts of LFOs in Minnesota at the community-level. Therefore, Minnesota, and Hennepin County specifically, may be a context where the LFOs of fines, fees, and restitution have muted effects at a community-wide level as compared to other states, or other counties in Minnesota, where courts are less responsive to defendants’ ability to pay. Another possibility is that the LFO effect reflects within-ZCTA gentrification practices (e.g., Hwang and Sampson 2014), where defendants with ample ability to pay are moving into more disadvantaged communities, thus diluting the overall impact of LFOs on disadvantage.

Finally, the analysis here tests the mediation hypothesis and finds that, consistent with vast neighborhood effects body of scholarship, concentrated disadvantage has a large and robust effect on levels of community violence. This provides evidence for the second link in the mediation hypothesis, as this completes the indirect link between community punishment concentrations and community violence. In terms of incarceration, the analysis revealed *bifurcating* direct and indirect effects, with a more substantial negative direct effect coupled with a countervailing positive effect via the indirect pathway through concentrated disadvantage. It should be noted that the analysis here is measuring relatively short-term effects of the punishment indicators (1-year lags). These findings about incarceration loads are consistent with research that show short-term benefits through incapacitation (Harding et al. 2017), but have more criminogenic long-term effects post-release after initial benefits of incapacitation wear off. Therefore, we may expect greater negative total effects over time, as the short-term incapacitation
effects wear off and the disadvantage induced by incarceration may compound itself and other sources of disadvantage over time, thereby resulting in more criminogenic longer-term effects through disadvantage accumulation (Kurlychek and Johnson 2019). In sum, these findings highlight how incarceration’s short-term benefits are partially washed out by its countervailing pathway that fortifies community disadvantage.

The fixed effect panel models and mediation analysis tell a slightly different story about probation loads, as probation was shown to have criminogenic direct and indirect effects on community violence. This is consistent with previous work in both Minnesota and California at the county-level (Santaularia, Larson, and Uggen 2021; Worrall et al. 2004), and recognizes community surveillance as a significant criminogenic force in the operations of communities – particularly in terms of violence. These findings give evidence towards the notion that it is not just incarceration that brings consequences for individuals and communities, and that the effects of different forms of punishment may differ. However, in contrast to incarceration, probation does not show any crime-reducing impacts either directly or indirectly. The indirect iatrogenic pathway of probation could be due, in part, to “system avoidance” (Brayne 2014) where individuals under state sanctioned surveillance are less likely to interact with key community social and economic institutions, thereby giving rise to higher levels of disadvantage. Further, the direct effects of probation estimated here, using hospital discharge records, are relatively uncontaminated by the mechanism of probation violations, as would be the case with more traditional measures of crime. Thus, the effect likely captures “new” violence as a result of probation, marking probation as an iatrogenic force for community crime and violence. It should also be noted that the areas high in probation are not just areas low in
incarceration, as evidenced by the strong correlations in Figure 8. However, the increase in violence as a result of changes in probation could also be a function of community surveillance, as injuries could be “detected” at a higher rate due to the increased community surveillance in some neighborhoods.

Finally, the analyses showed LFOs to have relatively smaller direct and indirect effects on community violence, with both pathways exhibiting a negative relationship to violence. While consistent in direction with deterrence perspectives, the negative relationships here also could be due to the considerations around ability to pay during the sentencing stage. As discussed above, the regime of monetary sanctions in Hennepin County may be “correcting” for the adverse effects of LFOs by making defendant ability to pay a forefront issue for county judges to take under consideration. Thus, the potential pathways by which LFOs may lead to further disadvantage and violence may be severed in Minneapolis.

Although the analysis here does find indirect effects of punishment via concentrated disadvantage, it should be noted that concentrated disadvantage, albeit a significant mediating force, is the only aspect of neighborhood structure investigated here. Put differently, concentrated disadvantage may not be the only neighborhood mechanism by which punishment impacts violence and crime. Thus, the direct effects, as specified in the models here, may operate through other facets of neighborhood ecology, or other forms of disadvantage not captured here (e.g., driver’s license suspensions). For example, punishment has also been shown to associate with other aspects of neighborhood structure. Harding et al. (2013) find that incarceration and conditions of post-prison release directly impact levels of community residential turnover, and
Massoglia et al. (2013) show that white prison returnees move to more disadvantaged communities post-release as compared to their respective pre-prison context. Therefore, it remains an open question to what extent punishment may structure other aspects of neighborhood structure, which in turn have implications for community violence. Future research could examine the other ways in which punishment may operate within this theoretical model, and highlight other iatrogenic pathways of punishment.

The empirical contributions here also speak to the theories of crime etiology that comprise ecological neighborhood theories. These theories have proven useful in describing the ecological factors influencing crime, as well as documenting the mediating forces of informal social control, networks, and collective efficacy that translate neighborhood structure into crime. However, these theories have overlooked the effects of political economy, public policy, and mechanisms of formal social control as key facets of community life. The results in this chapter suggest that punishment in its varied modalities can not only impact levels of crime and violence directly, but also indirectly by (re)structuring aspects of neighborhood ecology. I believe that including aspects of these “exogenous” forces within the ecological neighborhood model of crime will help improve the theorization of how neighborhood structure translates into crime, as well as provide a richer understanding how punitive policies can destabilize community life.

Alongside these theoretical implications, the empirical findings here also hold implications for future policy-making, particularly around the areas of punishment, policing, crime, and disadvantage. In particular, the results show that although incarceration may have certain short-term incapacitation effects, these public safety gains are diluted, suppressed, and partially negated by its penchant for inducing community
disadvantage, bringing the “incarceration ledger,” at least in terms of public safety, closer towards net zero. Further, this does not take into account the longer-term effects of punishment concentrations, which could increase over longer periods of exposure (e.g., Wodtke et al. 2011). More broadly, these findings highlight how policies that may help weaken the bond between punishment and disadvantage, such as ban-the-box and background check related policies (Agan and Starr 2018), record expungements, restrictions on record information access (Lageson 2020), or LFO debt forgiveness programs, may be productive avenues for combatting the iatrogenic effects of punishment on community violence. More strongly, a case could be made that punishment is spatially concentrated to such a degree that it has adverse consequences for disadvantage, and therefore diluting the concentration of punishment, particularly in the most disadvantaged places, would be fruitful areas for policy to consider (Wacquant 2009).

Although these analyses provide strong evidence for the effects of concentrated punishment on neighborhood structure and community violence, the study is not without its limitations. First, although the study here has robust controls through the fixed effects panel specifications, endogeneity in the form of omitted variables bias cannot be ruled out completely. Second, the nature of the available administrative data leaves a relatively short panel available to detect changes over-time, and a longer-term panel may be more amenable for looking at the longer-term effects of punishment on both disadvantage and violence. The data also limits the geographical unit of analysis to the ZCTA-level, within which substantial variability exists. Therefore, if more resolute address data on hospital discharges were available, the analysis could be taken to a lower level of geographical analysis (e.g., Census tract) where measures of violence and punishment may have more
variation amenable for analysis. Finally, the results here are valid for a specific
spatiotemporal context, Minneapolis 2011-2014. The extent to which the findings here
generalize to other times or contexts is a question for future research.

Despite the caveats mentioned above, this empirical dissertation chapter here
provides quantitative evidence of the iatrogenic link between punishment and crime, and
suggests that punishment can have criminogenic effects via neighborhood structure,
namely, concentrated disadvantage. Further, it has detailed the varied effects on violence
and disadvantage between the varying modalities of punishment, as well as the
bifurcating effect of incarceration. Broadly, this highlights that the effects of community
punishment regimes are not monolithic entities. Rather, they are varied systems of crime
reducing, crime inducing, and disadvantage promoting effects, all of which not only have
implications for crime but community well-being overall. The next chapter of this
dissertation turns to an empirical investigation of an ameliorative policy that has the
potential to suppress an iatrogenic effect of punishment on disadvantage: ban-the-box
policies in hiring decisions.
Chapter 3: Ban-The-Box, Employment, and Crime

Overview

This empirical chapter investigates a state policy with the potential to mitigate some of the iatrogenic effects of punishment on crime via employment: ban-the-box (BTB) legislation. While previous research has investigated the effects of these policies for individual employment, questions remain as to whether these laws impact macro-level employment and, distally, crime rates. This chapter leverages a uniquely constructed panel dataset of U.S. states from 1995-2020 and estimates staggered adoption difference-in-difference (DID) models to estimate the causal effect of BTB legislation on both employment and crime at the state-level. I estimate an average treatment effect of -1.7 for nonemployment, marking BTB laws as catalysts for state-level employment, but that these effects are moderated by both race and a state’s history of criminal record production. The results on crime are much weaker, suggesting that BTB legislation does not reach to significantly reduce state-level crime.
Introduction

The era of mass punishment – characterized by intense and concentrated loads of incarceration (Garland 2001), probation (Phelps 2010), and monetary sanctions (Harris 2016) – has received academic criticism for its inefficient crime reducing capabilities and collateral damage to both individuals and communities, particularly communities of color (Alexander 2012; Sharkey 2018a). Another important consequence of the mass incarceration and probation eras of U.S. punishment has been the proliferation of felony-level criminal records amongst the population. As of 2010, about 14.5 million Americans had a felony record but were no longer under any current correctional control. From 1980-2010, the share of American adults living in the community with a felony-record history increased by 3.8 percentage points up from 2.4 percent in 1980, a 260 percent increase (Shannon et al. 2017). Given the net cast by the institutions of criminal justice, as well as the racial disparities across each stage of the criminal justice process, the rise in felony records has been disproportionate by race and varies substantially between states (Shannon et al. 2017). The total U.S. adult population with felony records is about 8 percent, while the share of African American adults is about 23 percent, close to a quarter of the Black population in the U.S. This rate is even higher for African American men, where one-third have a felony record (Shannon et al. 2017). These trends in criminal record production, coupled with the vast research on collateral consequences (e.g., Uggen and Stewart 2015; Kirk and Wakefield 2018), have led scholars to conceptualize the criminal justice system as a key cog in the (re)production of American stratification (Wakefield and Uggen 2010). A central inequality-promoting mechanism
within this scholarship is the impacts of records within the formal labor market (e.g., Larson et al. 2022).

Both theory and empirical evidence suggest that criminal records reduce the chances of record holders of obtaining employment or progressing through various stages of the job interview process (e.g., Pager 2003), as well as suppress earnings relative to similarly situated peers (Western 2002). Despite the documented adverse effects of records on employment, a long history of sociological and criminological research shows that stable employment can be a key life course transition that promotes desistance from crime (Sampson and Laub 1993; Uggen 2000). Further, research has shown that the impacts of criminal record production register at the macro-level, with state-level felony history shares being linked to depressed state-level employment (Larson et al. 2022). Thus, criminal records may limit access to an integral social institution that promotes desistance from crime. Taken together, this scholarship suggests policies that alter information in regards to criminal records may not only lead to increased employment amongst record holders, but also decreased levels of crime.

These concerns have motivated the “ban-the-box” (BTB) movement, which involves states and other municipalities passing legislation that prevents employers from asking about criminal records until certain stages of the employment process (Vuolo et al. 2017). Advocates of BTB policies claim that record-based discrimination will be, perhaps partially, mitigated by BTB legislation, which could remove barriers to employment for record holders, and potentially serve as a public safety benefit. Therefore, BTB policies may be a potentially beneficial route to mitigate the harmful effects of criminal records in
the employment process, thus suppressing a potentially iatrogenic arm of punishment, at least within the arena of the labor market.

While previous scholarship on the effects of BTB legislation on employment have been promising in bolstering employment amongst those with criminal histories (e.g., Craigie 2020), other work has documented the racial heterogeneity in the effect of BTB and shown that it is not a remedy for all in the labor market. Specifically, BTB legislation has been shown to be detrimental for the employment prospects of Black and Hispanic applicants, as employers “statistically discriminate” in the midst of imperfect information and leverage their preexisting stereotypes of race and criminality as a proxy and substitute for criminal record information (Agan and Starr 2018; Doleac and Hansen 2020). This calls into question the efficacy of BTB laws, and their ability to help those who are more populous under the reach of the criminal legal system.

Despite this relatively recent research, questions remain as to the extent to which BTB policy influence macro-level economic and social well-being, as well as how the quality or scope (i.e., public and private BTB) moderate the influence the BTB legislation can have. Despite the theoretical and empirical connections between criminal records, employment, and crime, studies to date also have yet to consider the extent to which BTB legislation may ultimately impact the distal outcome of crime via the causal pathway of employment. Further, while studies have done work to isolate the populations or areas where BTB legislation may be most effective (e.g., Doleac and Hansen 2020, Shoag and Veuger 2021), no study to date has directly considered a state’s history of criminal record production and how that may impact the efficacy of BTB legislation in terms of employment and crime.
Thus, this empirical chapter sets out to investigate BTB legislation as a mechanism within the realm of state policy that has the potential, and manifest goals, of reducing the iatrogenic effects of records on state-level employment and crime. I ask three related research questions: 1) What are the macro-level effects of BTB legislation on both employment and crime at the state-level? 2) To what extent does the scope of a BTB law moderate its influence? And 3) To what extent does race and aggregate felony history alter the impacts of BTB? Thus, this third empirical chapter seeks fill these gaps in the literature by a) investigating the macro-level impacts on both employment and crime, with an eye towards not only the timing of law passage but also the quality of BTB adoption, and b) examine the potential racial and felony history heterogeneity in BTB laws’ effects by using a novel measure of felony-level criminal records (Shannon et al. 2017).

To answer these research questions, I construct a unique panel dataset of U.S. States from 1995-2020 and leverage the timing of BTB policy at the state level in a staggered adoption differences-in-differences design to estimate the causal impacts of BTB legislation on both employment and crime across a variety of different treatment and outcome specifications. In general, I find that BTB legislation offers an ameliorative policy route by which to combat the stigma of criminal records in the labor market, with BTB adoption being causally linked to a 1.7 percentage point reduction in overall nonemployment. This provides rigorous empirical evidence that state policy can, and has, reduced some of the barriers in the labor market for those with criminal records, thereby bolstering macro-level economic well-being in terms of employment. I also show that BTB laws greater in scope tend to be more efficacious. However, BTB laws are not a
panacea, as I find that BTB laws, while helping White employment rates, hinder Black employment rates, which is consistent with some previous research (e.g., Doleac and Hansen 2020). Finally, while BTB legislation may be a useful corrective in terms of overall employment, its effects are far weaker in the realm of crime.

I begin this chapter by describing some of the previous relevant scholarship on the connections between BTB legislation, employment, and crime, and move to discuss the construction of a unique panel dataset of U.S. States from 1995-2020. I then discuss the analytical strategy in terms of both the descriptive analysis, as well as the staggered adoption DID design and accompanying group-time average treatment effect estimation strategy (Callaway and Sant’Anna 2021a). Next, I walk through the findings of the various descriptive results and DID specifications, and end by discussing the findings and their policy implications, particularly in terms of BTB efficacy to mitigate the iatrogenic effects of records.

**Literature Review**

*Criminal Records and Employment*

Extensive scholarship on collateral consequences (Uggen and Stewart 2015; Kirk and Wakefield 2018) – consequences net of those part of the original sentence – has established the deleterious effects that the mark of a criminal record has in other social institutions such as employment (Pager 2003; Uggen et al. 2014), voting (Manza and Uggen 2008, Uggen et al. 2020), education (Stewart and Uggen 2020; Bernburg and Krohn 2003) and other aspects of social life. While the aftereffects of punishment may be wide-ranging, the large portion of social science scholarship has focused on the impacts of records on employment at both the individual and aggregate levels.
Two primary theoretical perspectives suggest that criminal records may have an adverse impact on prospects for employment. The stigma or market signal perspective suggests that the record acts as a signal to employers and can act as a form of statistical discrimination (Arrow 1973). Beyond employer discrimination, incarceration and other forms of punishment may diminish human capital (i.e., less education and job experience) amongst criminal justice impacted individuals (Western 2006; Raphael and Stoll 2013). Further, criminal justice impacted individuals have also been shown to have higher rates of health problems, such as chronic health issues (Schnittker and John 2007), infectious diseases (Massoglia 2008a), mental health (Sugie and Turney 2017), and lower self-reported health (Massoglia 2008b) (for a review, see Massoglia and Pridemore 2015), which may create further health-related barriers to employment post-punishment.

In general, survey-based studies find that incarceration is associated with a 10-20 percentage point reduction in employment (e.g., Apel and Sweeten 2010), whereas studies using administrative data find generally weaker or no effects (e.g., Harding et al. 2018). It should be noted that these studies generally measure the effects of incarceration on employment outcomes, as opposed to a criminal record.

Another set of studies isolate the impacts of criminal records by leveraging experimental audit designs to examine the “credentialing” effect of criminal records on job callbacks, and generally find that the presence of a criminal record reduces employment likelihoods or job interview “callbacks,” and this effect is conditioned by the applicant’s race (Pager 2003; Pager, Bonikowski, and Western 2009; Uggen et al. 2014). Pager (2003), in a classic study leveraging an experimental field audit design, sent testers to in-person job applications in Milwaukee and experimentally manipulated the race and
criminal record of applicants. Among White applicants, the presence of a criminal record decreased the likelihood of a callback by 50% as compared to the non-record White control. Further, Pager (2003) also found an effect of race, with the Black testers without records less likely to receive a job callback than the White testers without records. However, race also interacts with the presence of a criminal record, with the criminal record penalty for Black testers being 40% larger than that of the White tester.

Pager et al. (2009), in a follow-up audit in New York City, used two teams of White, Black, and Latino testers and found that White testers received more callbacks than either Latino or Black testers, and Latino testers were preferred over Black applicants. However, when comparing Black and Latino testers to a White tester with a criminal record, they found no statistically significant differences in callback rates indicating that “while ex-offenders are disadvantaged in the labor market relative to applicants with no criminal background, the stigma of a felony conviction appears to be no greater than that of minority status (10).” While Pager (2003) and Pager et al. (2009) signaled a criminal record with a drug possession felony, Uggen et al. (2014) examine “the edge of stigma” and signal a disorderly conduct arrest that did not lead to a conviction. They find roughly a four-percentage point difference between the treatment conditions for both White and Black testers and find a statistically significant effect of misdemeanor arrest on callback rates once adjusting for contact with the hiring entity. In their Minnesota audit, Black testers were called back 33% less often than the white testers, which is a smaller race effect than reported in Pager (2003). Finally, they also find that the misdemeanor effect interacts with workplace diversity, with the misdemeanor arrest having little effect at establishments with greater workforce diversity.
The relationship between criminal records and employment has also been corroborated at the state-level. Using fixed effects panel models, Larson et al. (2022) find that a one percentage point increase in the proportion of adults with a felony-level criminal record increases non-employment rate by .3 percentage points. Further, they find that the deleterious effects of felony record production are greater amongst White individuals, which they interpret to possibly due to the overriding stigma of race in employment interactions. In sum, a wide breadth of rigorous scholarship has shown the relationship between criminal records and depressed employment at both the individual and aggregate levels.

**Employment and Crime**

While criminal records have shown to significantly impact employment across multiple levels of analysis, scholarship on the second path in this causal chain regards the linkages between employment and crime. Early theorizing in economic models of crime have suggested that more plentiful and stable employment opportunities reduce crime, as those employed experience less utility from illegal behavior in the presence of gains from employment (Becker 1968). Going beyond rational choice models of crime, sociological theories of crime all carry implications for the employment-crime link. Social control theories (Hirschi 1969; Sampson and Laub 1993) see employment as a key element of social capital that bonds individuals to the conventional social order, with these bonds varying in intensity over the life course. Routine activities theories contend that employment shifts the structure of individuals’ routine actives, with more time spent at work (e.g., Osgood et al. 1996). Learning theories, such as differential association (Sutherland 1947; Matsueda 1988) would hold that employment can alter social
networks, and the resulting definitions, attitudes, and reinforcements surrounding
criminal behavior. Strain theories (Merton 1938; Agnew 1992) suggest that success in the
labor market may prevent disjunctures between goals and means, and the resulting strain
that ensues from such disconnections, thereby decreasing crime. Finally, social
interaction and labeling theories of crime may predict that employment brings changes in
reference groups and non-criminal opportunities, bringing alterations of self-conceptions
(Matsueda 1992).

A wide swath of empirical scholarship has shown crime to be negatively
correlated with employment at the individual-level. Studies using survey methods find
consistent negative relationship between employment and crime (e.g., Apel et al. 2008).
Life course studies also show that employment can serve as a “turning point” in the lives
of individuals and promote desistance from crime (Sampson and Laub 1993; Laub and
Sampson 1993). Research in the reentry context has also shown that employment can
reduce crime, recidivism, and illegal earnings (Uggen 2000; Uggen and Thompson 2003;
Yang 2017), and that the subjective experience and commitment to work, opposed to
other work characteristics, are integral for this relationship (Apel and Horney 2017). This
link has also been examined at more macro-levels of analysis. Cantor and Land (1985)
discuss two competing mechanisms within the unemployment-crime link, with increases
in crime due to increases in strain and decreases in informal social control offset, perhaps
partially, by decreases in criminal opportunities due to alterations in routine activities.
While studies have shown that there is empirical validity to each of these
mechanisms across multiple levels of analysis (e.g., Phillips and Land 2012), studies have
generally found negative, but primarily weak, relationships between employment and
crime at the aggregate level (e.g., Levitt 2001; Arvanites and Defina 2012), and weaker relationships with violent crime as compared to property crime (Raphael and Winter-Ebmer 2001). Overall, these studies are relatively consistent in that employment is negatively associated to crime, with a corollary being that punishment may exacerbate crime through its indirect effect on employment.

The existing theory and empirical evidence suggest that criminal records have an indirect effect on crime through the mediator of employment. Thus, if state policy is able to “flip the lever” of employment at the macro-level, previous evidence on the employment-crime link suggests that the passage of legislation such as BTB could leverage the mediating force of employment to pattern crime rates. However, I stress that previous research suggests that this effect may possibly be small, as research finds that the employment-crime link is relatively weak (Levitt 2001; Arvanites and DeFina 2006), and employment shows relatively weaker relationships with violent crime as compared to property crime (Raphael and Winter-Ebmer 2001).

Ban-the-Box Legislation

While a robust literature establishes a relationship between criminal records and both employment and crime, relatively new public policy concerning BTB legislation demarcate attempts via policy to counteract the effects a criminal record has on employment. A common practice in the United States labor market is to screen applicants on the basis of a criminal record and conduct background checks (Bushway 2004; Raphael 2006), with rationale for the practice including fears of liability, lower-levels of educational attainment and work experience, and distrust of individuals holding criminal records (Raphael 2021). Although survey estimates vary, an estimated 60%-75% of
employers consider criminal histories when making hiring decisions (Raphael 2021), and organizational variation exists in the ways in which employers use criminal history information (Lageson et al. 2015). Scholarship finds considerable diversity in the type and breadth of criminal record questions asked on applications, as well as significant variation in the probability of asking a criminal record question amongst the axes of workplace diversity, labor type, and neighborhood disadvantage (Vuolo et al. 2014). Further, adoption rates of criminal background check laws have been shown to be driven by a states’ racial composition and racial economic threat, as opposed to levels of crime, with larger shares of Black residents, particularly in places with higher White unemployment, being associated with higher rates of background check laws (McElhatten 2022). This provides evidence that “stereotypes of Black criminality shape lawmakers’ implicit judgments to produce racially patterned policy outcomes (McElhatten 2022:1079).”

Although sociological theory suggests that jobs are important in the reentry process, studies examining reentry-based employment programming generally find relatively weak effects (Bushway and Apel 2012), or effects concentrated among subgroups (Uggen 2000; Uggen and Shannon 2014). Despite these findings, states have begun to go beyond correctional programming in the reentry stage and attempt to remove barriers in the employment process by limiting the accessibility, amount, and timing of criminal history directly in the labor market. A recent literature base, primarily based in economics, examines the proliferation of BTB legislation as well as the extent to which BTB policies alleviate the barrier placed upon record holders in the labor market.
BTB policies attempt to mitigate these discriminatory effects by banning criminal record questions from applications as well as pushing background checks off until a later employment stage. BTB policies can vary greatly in the details and minutiae, but generally follow a consistent pattern across states and fall into three groups with varying scope: 1) state level employment, 2) all public employment (including state and municipal level employment, and sometime private employment with governmental contracts), and 3) full private employment bans (Avery and Lu 2021; Doleac and Hansen 2020). BTB policies have begun to garner political traction in the U.S. with 35 states adopting some form of BTB legislation as of 2020 (Avery and Lu 2021), with an estimated 80% of Americans living in a jurisdiction with some variety of BTB legislation (Avery and Lu 2021). Doleac and Hansen (2017) find that cities with greater shares of Black residents, as well as higher rates of college degrees are more likely to pass a BTB law. Further, they show that BTB adoption is not responsive to demographic shifts or migration patterns between labor markets.

Recent research on the effects of “ban the box” policy implementation on employment has found that these policies may reduce the impacts of records for white applicants, but they may worsen employment prospects for Black applicants (Raphael 2021). Craigie (2020) uses a DID design and finds that jurisdictions that adopted BTB legislation showed increases in the likelihood of public sector employment amongst those with criminal records at the individual-level with no differences in effect magnitude amongst Black and White individuals. Similarly, Shoag and Veuger (2021), also using a DID design, find an increase in employment in “high-crime” census tracts. Studies using administrative data also assess the impacts of BTB laws, finding that employment and
earnings decline slightly for individuals with criminal history (Jackson & Zhao 2017).

Similarly, Seattle’s 2013 BTB ordinance did not increase the relative employment or earnings amongst those with criminal histories in Seattle as compared to those living elsewhere in Washington (Rose 2021).

Despite these promising, yet mixed, findings, other scholarship indicates that BTB may not improve the employment prospects for all racial groups, and it may in fact be detrimental to employment for Black individuals without criminal records. Early inquiry into the automation of state-level criminal history databases finds that the greater accessibility of criminal records is associated with a higher ratio of black wages relative to white wages (Bushway 2004), suggesting that the withholding criminal history information may worsen the economic lives of Black individuals in the aggregate. In a field experiment of New Jersey and New York, Agan and Starr (2018) submitted fictitious job applicants, experimentally manipulating criminal record and race of the applicant, and found that after BTB passage, the gap between White to Black callbacks increased sixfold, suggesting that employers use race as a proxy for criminal history in employment decisions. Doleac and Hansen (2020), exploiting variation in the adoption of both state and local BTB policies, also find negative effects of BTB legislation on employment likelihoods amongst young Black and Hispanic men at the individual-level. Further, they show not only that these effects increase over time, but that the effects vary with the local labor market conditions, where places with a higher proportion of young Black and Hispanic men exhibit less harmful effects. These race findings are consistent with survey-based studies that find that employers who conduct criminal background
checks are more likely to hire Black applicants, particularly amongst those most averse to hiring applicants with criminal histories (Holzer et al. 2006).

These patterns amongst racial groups are also corroborated by other types of screening bans, such as drug testing (Wozniak 2015) and credit checks (Bartik and Nelson 2019). In contrast, a Minnesota field audit found nonsignificant differences in callback rates amongst Black applicants with no records between employers with or without criminal record questions (although the gradient between question types was consistent with statistical discrimination) (Vuolo et al. 2014). Given the potential heterogeneous effects of race observed in previous scholarship, the current analysis also estimates DID specifications conditioned on racial group.

While not a direct corollary, it could be that state-level felony history works in a similar fashion to that of race: places with a high proportion of criminal records may reduce the impact of BTB legislation, as criminal records are more prevalent in the local labor market. Put differently, employers structurally are given more latitude to exclude individuals with criminal records when these applicants are plentiful relative to the number of available positions. However, some empirical evidence points towards the opposite direction. Shoag and Veuger (2021), mentioned previously, finding relative increases in employment amongst “high-crime” neighborhoods as compared to “low-crime” neighborhoods. Despite the usage of crime-rates as a proxy for the prevalence of criminal records, this study suggests that BTB policies may be more efficacious in area with high levels of punishment and criminal record production. Therefore, this empirical chapter will provide a more direct test of this hypothesis by examining the treatment
effects of BTB legislation amongst states with varying levels of estimated felony history shares.

Data and Measures

This empirical chapter utilizes a unique panel dataset of U.S. States from 1995-2020 constructed from a variety of sources to investigate the relationships between BTB legislation, employment, and crime. These data consist of spatial shapefiles from the Census Bureau’s Cartographic Boundary Shapefiles (Walker 2022), sociodemographic and economic data from the Current Population Survey, policy measures from the University of Kentucky Center for Poverty Research’s National Welfare Data, estimates of criminal record production from life table estimates (Shannon et. al. 2017), and data on BTB legislation from the National Employment Law Project.

National Employment Law Project Data

Data on BTB legislation by state was collected from the National Employment Law Project (Avery and Lu 2021). Each state’s policy was coded for the year a policy went into effect, as well as the scope of the law that was passed. The scope pertains to the breadth of employment that the law that bans the box and background checks, including state employment, full public employment (e.g., state and municipalities), and private employment, as well as combinations therein. As of 2020, 34 states had passed some form of BTB legislation, and the BTB policies enacted can be found in Table 1.

Generally, states follow a progression from starting with a ban of more limited scope (e.g., state employment only) and passing broader legislation after the passage of the first law (e.g., California). However, some states (e.g., Oregon) pass full public and private bans with their first BTB law. Given this pattern of evolution of BTB legislation,
I create three primary treatment indicators to measure the impact of BTB. First, I define treatment as the passage of a state’s first BTB law of any scope (i.e., comparisons between any BTB, regardless of evolution, and no BTB), and this serves as the treatment in the primary specifications below. Second, treatment is defined as the passage of at least a full public ban, which compares states that have full public bans or better to states with no BTB or only BTB that applies to state employment. Treatment is lastly defined as the passage of a full public and private ban, the most complete form of BTB legislation observed as of the end of the study period. A state-year is considered treated in this indicator if it: a) passed a full public/private ban combination law, b) passed a private law after already having a full public ban in place. Thus, these three treatment thresholds serve as the key treatment indicators in the empirical analysis below. The BTB laws included in this analysis can be found in Table 12.

23 It should be noted that individual states’ BTB legislation contain exemptions for certain jobs, fields of employment, size of employers, or certain charges (e.g., “crimes of moral turpitude”) making laws unique in some regards. Further, some states will differ in the timing in which they allow employers to eventually ask or seek out criminal background related information (e.g., after first interview, after offer of employment, etc.). However, these qualities of the law are likely to only modestly influence the impacts laws have on outcomes, at least compared to the scope of the law.
Table 12: BTB State-Level Legislation

<table>
<thead>
<tr>
<th>State</th>
<th>Effective Year</th>
<th>Legislation</th>
<th>Legal Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona</td>
<td>2017</td>
<td>Executive Order 2017-07</td>
<td>Public</td>
</tr>
<tr>
<td>California</td>
<td>2010</td>
<td>NA</td>
<td>State</td>
</tr>
<tr>
<td>California</td>
<td>2013</td>
<td>AB 218</td>
<td>Public</td>
</tr>
<tr>
<td>California</td>
<td>2017</td>
<td>AB 1008 California Fair Chance Act</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Colorado</td>
<td>2019</td>
<td>HB 19-1025</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Connecticut</td>
<td>2017</td>
<td>HB 5237</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Delaware</td>
<td>2014</td>
<td>HB 167</td>
<td>Public</td>
</tr>
<tr>
<td>Georgia</td>
<td>2015</td>
<td>NA</td>
<td>State</td>
</tr>
<tr>
<td>Hawaii</td>
<td>1998</td>
<td>HR 3638</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Illinois</td>
<td>2013</td>
<td>Executive Order 1(CM100)</td>
<td>State</td>
</tr>
<tr>
<td>Illinois</td>
<td>2015</td>
<td>HB 5701</td>
<td>Private</td>
</tr>
<tr>
<td>Indiana</td>
<td>2017</td>
<td>Executive Order 17-15</td>
<td>State</td>
</tr>
<tr>
<td>Kansas</td>
<td>2018</td>
<td>Executive Order 18-12</td>
<td>State</td>
</tr>
<tr>
<td>Kentucky</td>
<td>2017</td>
<td>Executive Order</td>
<td>State</td>
</tr>
<tr>
<td>Louisiana</td>
<td>2016</td>
<td>Act 398 (HB 266)</td>
<td>State</td>
</tr>
<tr>
<td>Maine</td>
<td>2019</td>
<td>L.D. 170</td>
<td>State</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>2010</td>
<td>Chapter 256 of the Acts of 2010 (Senate Bill 2583)</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Michigan</td>
<td>2018</td>
<td>Executive Directive 2018-4</td>
<td>State</td>
</tr>
<tr>
<td>Minnesota</td>
<td>2009</td>
<td>HF 1301</td>
<td>Public</td>
</tr>
<tr>
<td>Minnesota</td>
<td>2014</td>
<td>SF 523</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Missouri</td>
<td>2016</td>
<td>Executive Order 16-04</td>
<td>Public</td>
</tr>
<tr>
<td>Nebraska</td>
<td>2014</td>
<td>LB 907</td>
<td>Public</td>
</tr>
<tr>
<td>Nevada</td>
<td>2018</td>
<td>Assembly Bill 384</td>
<td>Public</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>2020</td>
<td>House Bill 253</td>
<td>State</td>
</tr>
<tr>
<td>New Jersey</td>
<td>2015</td>
<td>The Opportunity to Compete Act (A1999)</td>
<td>Public and Private</td>
</tr>
<tr>
<td>New Mexico</td>
<td>2010</td>
<td>S.B. 254</td>
<td>Public</td>
</tr>
<tr>
<td>New Mexico</td>
<td>2019</td>
<td>S.B. 96</td>
<td>Private</td>
</tr>
<tr>
<td>New York</td>
<td>2015</td>
<td>NA</td>
<td>State</td>
</tr>
<tr>
<td>North Carolina</td>
<td>2020</td>
<td>Executive Order 158</td>
<td>State</td>
</tr>
<tr>
<td>North Dakota</td>
<td>2019</td>
<td>HB 1282</td>
<td>Public</td>
</tr>
<tr>
<td>Ohio</td>
<td>2015</td>
<td>HR-29 and HB 56</td>
<td>Public</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>2016</td>
<td>Executive Order 2016-03</td>
<td>State</td>
</tr>
<tr>
<td>Oregon</td>
<td>2015</td>
<td>HB 3025</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>2018</td>
<td>Administrative Policy HR-TM001</td>
<td>State</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>2014</td>
<td>HB 5507</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Tennessee</td>
<td>2016</td>
<td>SB 2440</td>
<td>State</td>
</tr>
<tr>
<td>Utah</td>
<td>2017</td>
<td>HB 156</td>
<td>Public</td>
</tr>
<tr>
<td>Vermont</td>
<td>2015</td>
<td>Executive Order</td>
<td>State</td>
</tr>
<tr>
<td>Vermont</td>
<td>2017</td>
<td>HB 261</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Virginia</td>
<td>2015</td>
<td>Executive Order 41</td>
<td>State</td>
</tr>
<tr>
<td>Virginia</td>
<td>2020</td>
<td>HB 757</td>
<td>Public</td>
</tr>
<tr>
<td>Washington</td>
<td>2018</td>
<td>HB 1298</td>
<td>Public and Private</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>2016</td>
<td>Assembly Bill 373</td>
<td>State</td>
</tr>
</tbody>
</table>

Current Population Survey Data

After the construction of state-year indicators of BTB treatment, I merge onto this panel data state-year aggregated measures from the Current Population Survey (CPS)
Basic Monthly data\textsuperscript{24}, as well as the Annual Social and Economic Supplement of the CPS (ASEC). The CPS is a multistage, stratified probability-selected monthly survey of \textasciitilde60,000 households conducted by the U.S. Census Bureau, with microdata compiled by the IPUMS USA project, that represents one of the leading data sources for labor force statistics and other demographic data (Flood et al. 2021). The primary sampling stage involves 852 sample areas nested within States, within which a sample of households is drawn. Every March, the ASEC supplement is conducted on \textasciitilde70,000 households where more detailed economic and social data is captured. The ASEC sample primarily consists of the March CPS Basic Monthly sample, as well as additional CPS households.

From the CPS Basic Monthly micro data, I construct the focal outcome of interest: the nonemployment share for the prime age (25-54) civilian, non-institutionalized adult population. The nonemployment rate is equals the percentage of people unemployed or not in the labor force as a ratio of the entire population. The nonemployment share is also the complement of the employment-to-population ratio, and the nonemployment share is chosen over other measures of employment because it will capture two forms of potential BTB alleviation or exacerbation: the stigma or statistical discrimination of criminal records for those participating in the labor market, and the record effects that lead to individuals to withdraw from the formal labor market altogether (Larson et al. 2022; Apel and Sweeten 2010). I also aggregate population age shares (16-25, 26-35, 36-45, 46-55, 56-65, 65+), prime-age marriage, bachelor’s degree, unemployment and poverty rates from the Basic Monthly files, as well as disability rates.

\textsuperscript{24} The Basic Monthly microdata weights are weighted to the entire U.S. population for each given month, and to aggregate to a yearly estimate using these weights I adjust each monthly weight by the ratio of the monthly sample to the pooled yearly sample, so that the weights for each case add up to a yearly population estimate.
from the ASEC supplement, which serve as key covariates and have been shown to associate with nonemployment shares (Aaronson et al. 2014; Austin et al. 2018; Larson et al. 2021). I also construct race-specific measures of each of these CPS measures, which are used in the race-subgroup models described below. These measures are used in the difference-in-difference models to control for any contemporaneous economic, social, and demographic changes with the passage of BTB legislation.\textsuperscript{25}

\textit{UKCPR National Welfare Data}

To control for aspects of social welfare and labor policy, I construct time-varying measures of state policy from the University of Kentucky Center for Poverty Research’s National Welfare Data (University of Kentucky Center for Poverty Research 2022). Following Larson et al. 2021, I create measures of the effective minimum wage (the maximum of the state or federal minimum wage in each state-year), and the mean Assistance to Families with Dependent Children (AFCD), Supplemental Nutrition Assistance Program, and/or Temporary Assistance to Needy Families (TANF) maximum benefit between 2-person, 3-person, and 4-person families. These measures serve as key control variables in the difference-in-difference models to exclude the possibility of contemporaneous social welfare changes alongside BTB policy enactment.

\textit{Federal Bureau of Investigation Uniform Crime Reports}

The second focal outcome, the crime rate, is sourced from the FBI’s Uniform Crime Reports program data. Specifically, I use the Offenses Known and Clearances by

\textsuperscript{25} Following previous scholarship (Aaronson et al. 2014; Larson et al. 2021), I also calculate 1–3-year lags of the unemployment rate to control for business cycle fluctuation. The inclusion of these covariates in the DID specifications below remove the earliest time periods (1995-1998) from the analysis due to listwise deletion in the estimation, which removes Hawaii’s 1998 BTB law from the treatment pool.
Arrest data, which comprises counts of index crimes – homicide, rape, robbery, aggravated assault, burglary, theft, motor vehicle theft, and arson - known or reported to police (Kaplan 2021a). These data also follow the Hierarchy Rule, in which only the most serious crime is recorded per individual incident, resulting in a modest undercounting of reported crime (Kaplan 2021a). Arsons, while considered an index (or “Part I” crime), are not included in this data and are therefore sourced from the FBI’s Arson dataset.

UCR data does not include crimes not reported or captured by law enforcement agencies, and it also relies on voluntary reporting by law enforcement agencies, which can result in no or incomplete data for some jurisdictions (Kaplan 2021a). Given this limitation in the numerator, using a full population estimate, such as the CPS estimated population, as the denominator in the rate calculation would result in suppressed crime rates, as places would be counted in the population denominator that are not included in the numerator. To avoid this issue, I aggregate each reporting agency’s primary jurisdictional population to the state level, which represents a population denominator of reporting agencies in each state. I use previously concatenated UCR Offenses Known data from 1960-2020 (Kaplan 2021b), to construct the index crime rate, which is expressed as the ratio of reported index crimes to the reporting jurisdiction state population multiplied by 100,000.

Felony History Shares

To evaluate if BTB legislation has differential effect on employment in states with greater criminal record production, I leverage life-table estimates of the adult (18+) population living in the community with a felony record but no longer under supervision (Shannon et al. 2017). Specifically, these estimates were calculated by taking annual
coHORTS OF PRISON RELEASE AND FELONY PROBATION ENTRIES FROM EACH YEAR, AND ADJUSTING EACH
COHORT FOR MORTALITY, RECIDIVISM, MOBILITY, AND DEPORTATION IN EACH SUBSEQUENT YEAR. THESE
ESTIMATES OFFER A COMPREHENSIVE, AND PRECISE, VIEW OF THE SHARE OF INDIVIDUALS WHO ARE AT
RISK OF SUFFERING LABOR MARKET CONSEQUENCES DUE TO A FELONY RECORD. THESE ESTIMATES ARE
ONLY AVAILABLE UNTIL THE YEAR 2010, AND THEREFORE I USE THE MOST RECENT AVAILABLE YEAR
(2010) TO MEASURE A STATE’S FELONY HISTORY SHARE. I CREATE A BINARY INDICATOR OF ABOVE AND
BELOW MEAN STATE FELONY HISTORY SHARES, AND ESTIMATE SPLIT SAMPLE DID SPECIFICATIONS TO
EXAMINE THE TREATMENT EFFECT HETEROGENEITY AMONGST VARYING STATE CRIMINAL RECORD
PREVALENCE.

BUREAU OF JUSTICE STATISTICS – NATIONAL PRISONER STATISTICS SURVEY

PREVIOUS RESEARCH HAS SHOWN THAT INCARCERATION RATES HAVE A MODEST RELATIONSHIP TO
CRIME RATES (E.G., LEVITT 1996), AND THEREFORE I LEVERAGE THE BUREAU OF JUSTICE STATISTICS
NATIONAL PRISONER STATISTICS SURVEY DATA TO ACCOUNT FOR TIME-VARYING STATE DIFFERENCES IN
INCARCERATION IN THE CRIME RATE DIFFERENCE-IN-DIFFERENCE SPECIFICATIONS. I AGGREGATE THE
TOTAL NUMBER OF INCARCERATED INDIVIDUALS AS OF YEAR-END IN EACH STATE-YEAR, AND I CALCULATE
THE INCARCERATION RATE AS THE SUMMED INCARCERATED PERSONS AS A RATIO OF THE CPS TOTAL
POPULATION (16+) ESTIMATE, MULTIPLIED BY 100,000. THIS RESULTS IN A RATE OF INCARCERATED
PERSONS PER 100,000 IN EACH STATE-YEAR. THE BJS HAVE YET TO RELEASE YEAR END
CRIME RATE SPECIFICATION, WHICH USES THE INCARCERATION RATE AS A KEY COVARIATE, UTILIZE A
STUDY WINDOW OF 1995-2019. DESCRIPTIVE STATISTICS FOR ALL VARIABLES IN THE ANALYSIS ARE
INCLUDED IN TABLE 13.
Table 13: Descriptive Statistics for All Variables

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTB-State</td>
<td>1,300</td>
<td>0.065</td>
<td>0.246</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BTB-Public</td>
<td>1,300</td>
<td>0.064</td>
<td>0.226</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BTB-Private</td>
<td>1,300</td>
<td>0.066</td>
<td>0.078</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BTB-Public/Private</td>
<td>1,300</td>
<td>0.061</td>
<td>0.239</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Prime: Nonemployment Share-Black</td>
<td>1,294</td>
<td>26.150</td>
<td>6.972</td>
<td>2.019</td>
<td>26.399</td>
<td>63.938</td>
</tr>
<tr>
<td>Total Index Crime Rate</td>
<td>1,300</td>
<td>3,305.971</td>
<td>1,158.407</td>
<td>0.000</td>
<td>3,163.293</td>
<td>7,981.526</td>
</tr>
<tr>
<td>Population Share 15-25</td>
<td>1,300</td>
<td>17.540</td>
<td>1.723</td>
<td>12.421</td>
<td>17.445</td>
<td>27.345</td>
</tr>
<tr>
<td>Population Share 25-35</td>
<td>1,300</td>
<td>17.387</td>
<td>1.855</td>
<td>12.554</td>
<td>17.166</td>
<td>24.088</td>
</tr>
<tr>
<td>Population Share 35-45</td>
<td>1,300</td>
<td>18.050</td>
<td>2.610</td>
<td>12.890</td>
<td>17.718</td>
<td>29.017</td>
</tr>
<tr>
<td>Population Share 45-55</td>
<td>1,300</td>
<td>17.286</td>
<td>1.636</td>
<td>12.423</td>
<td>17.314</td>
<td>23.033</td>
</tr>
<tr>
<td>Population Share 65+</td>
<td>1,300</td>
<td>16.015</td>
<td>2.586</td>
<td>5.637</td>
<td>15.755</td>
<td>24.854</td>
</tr>
<tr>
<td>Population Share 45-55-White</td>
<td>1,300</td>
<td>17.533</td>
<td>1.735</td>
<td>12.654</td>
<td>17.164</td>
<td>24.933</td>
</tr>
<tr>
<td>Population Share 65+</td>
<td>1,300</td>
<td>17.094</td>
<td>2.921</td>
<td>5.024</td>
<td>16.894</td>
<td>26.873</td>
</tr>
<tr>
<td>Population Share 15-25-Black</td>
<td>1,296</td>
<td>23.178</td>
<td>6.354</td>
<td>0.000</td>
<td>22.208</td>
<td>71.286</td>
</tr>
<tr>
<td>Population Share 35-45-Black</td>
<td>1,294</td>
<td>19.722</td>
<td>5.463</td>
<td>2.798</td>
<td>19.268</td>
<td>84.294</td>
</tr>
<tr>
<td>Population Share 45-55-Black</td>
<td>1,286</td>
<td>15.948</td>
<td>4.154</td>
<td>0.875</td>
<td>16.215</td>
<td>55.257</td>
</tr>
<tr>
<td>Population Share 55-65-Black</td>
<td>1,264</td>
<td>10.991</td>
<td>4.234</td>
<td>0.680</td>
<td>10.827</td>
<td>53.786</td>
</tr>
<tr>
<td>Population Share 65+</td>
<td>1,249</td>
<td>7.936</td>
<td>3.612</td>
<td>0.373</td>
<td>10.105</td>
<td>24.171</td>
</tr>
<tr>
<td>Prime Marriage Rate</td>
<td>1,300</td>
<td>62.592</td>
<td>4.874</td>
<td>49.531</td>
<td>62.393</td>
<td>77.980</td>
</tr>
<tr>
<td>Prime Marriage Rate-White</td>
<td>1,300</td>
<td>65.297</td>
<td>4.978</td>
<td>51.587</td>
<td>65.074</td>
<td>78.199</td>
</tr>
<tr>
<td>Prime Marriage Rate-Black</td>
<td>1,297</td>
<td>43.759</td>
<td>10.034</td>
<td>5.897</td>
<td>42.735</td>
<td>100.000</td>
</tr>
<tr>
<td>Prime Degree Rate</td>
<td>1,300</td>
<td>31.138</td>
<td>6.621</td>
<td>18.693</td>
<td>30.207</td>
<td>56.591</td>
</tr>
<tr>
<td>Prime Degree Rate-White</td>
<td>1,300</td>
<td>32.267</td>
<td>6.681</td>
<td>19.503</td>
<td>31.113</td>
<td>57.512</td>
</tr>
<tr>
<td>Prime Degree Rate-Black</td>
<td>1,292</td>
<td>22.394</td>
<td>8.683</td>
<td>2.138</td>
<td>20.792</td>
<td>100.000</td>
</tr>
<tr>
<td>Prime Disability Rate</td>
<td>1,300</td>
<td>7.712</td>
<td>2.682</td>
<td>3.349</td>
<td>7.125</td>
<td>29.231</td>
</tr>
<tr>
<td>Prime Disability Rate-White</td>
<td>1,300</td>
<td>7.237</td>
<td>2.590</td>
<td>3.030</td>
<td>6.710</td>
<td>28.849</td>
</tr>
<tr>
<td>Prime Disability Rate-Black</td>
<td>1,300</td>
<td>10.888</td>
<td>6.860</td>
<td>0.000</td>
<td>10.587</td>
<td>64.931</td>
</tr>
<tr>
<td>Prime Poverty Rate</td>
<td>1,300</td>
<td>10.680</td>
<td>4.338</td>
<td>2.504</td>
<td>9.640</td>
<td>47.907</td>
</tr>
<tr>
<td>Prime Poverty Rate-White</td>
<td>1,300</td>
<td>9.068</td>
<td>3.930</td>
<td>2.636</td>
<td>8.240</td>
<td>43.844</td>
</tr>
<tr>
<td>Prime Poverty Rate-Black</td>
<td>1,300</td>
<td>19.394</td>
<td>12.013</td>
<td>0.000</td>
<td>18.375</td>
<td>104.644</td>
</tr>
<tr>
<td>Effective Minimum Wage</td>
<td>1,300</td>
<td>6.571</td>
<td>1.645</td>
<td>4.250</td>
<td>6.750</td>
<td>13.500</td>
</tr>
<tr>
<td>AFDC/TANF/SNAP Maximum</td>
<td>1,300</td>
<td>846.762</td>
<td>201.858</td>
<td>429.667</td>
<td>808.867</td>
<td>1,541.667</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>1,300</td>
<td>5.473</td>
<td>1.930</td>
<td>2.217</td>
<td>5.080</td>
<td>14.523</td>
</tr>
<tr>
<td>T_{1} Unemployment Rate</td>
<td>1,250</td>
<td>5.398</td>
<td>1.977</td>
<td>2.217</td>
<td>5.015</td>
<td>14.523</td>
</tr>
<tr>
<td>T_{2} Unemployment Rate</td>
<td>1,200</td>
<td>5.471</td>
<td>1.984</td>
<td>2.217</td>
<td>5.089</td>
<td>14.523</td>
</tr>
<tr>
<td>T_{3} Unemployment Rate</td>
<td>1,150</td>
<td>5.542</td>
<td>1.896</td>
<td>2.217</td>
<td>5.164</td>
<td>14.232</td>
</tr>
<tr>
<td>Incarceration Rate</td>
<td>1,250</td>
<td>54.0123</td>
<td>215.588</td>
<td>129.575</td>
<td>513.496</td>
<td>1,199.470</td>
</tr>
</tbody>
</table>

**Analytical Strategy**

*Descriptive Analyses of BTB, Employment, and Crime*

The first part of the analysis involves descriptive and bivariate analyses of the status of BTB law passage, the spatiotemporal patterns of nonemployment and crime.
rates, and the bivariate relationships between BTB law passage and each outcome measure. First, I create a choropleth map of the status of BTB legislation by state as of the end of the study period, 2020. The shapefile for the United States was obtained from the Census Bureau API via the ‘tidycensus’ package in R (Walker and Herman, 2022), and the mapping uses the USA Contiguous Albers Equal Area Conic (ESRI:102003) projection (Walker 2022). Hawaii and Alaska are shifted and rescaled before being assigned to the coordinate reference system in order to display them alongside the contiguous US.

To examine the event history of BTB law passage in the United States, I calculate Kaplan-Meier survival estimates to examine the survival distributions of states that have yet to pass a BTB law. The Kaplan-Meier estimator estimates the proportion of states that have yet to pass a BTB law as of a given year, and is expressed as:

\[ S(t) = \prod_{i: t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \]

which effectively conditions each successive survival probability by the probability of surviving to a given point in time, \( t \) (Kaplan-Meier 1958). Following the coding scheme of the treatment indicators as discussed above, I calculate Kaplan-Meier estimates for three separate events: the passage of a state’s first BTB law of any scope, the passage of a BTB law that at least includes a full public ban, and the passage of a BTB law that includes a full public/private ban. Each Kaplan-Meier survival curve is estimated using the ‘survival’ package in R (Therneau 2022).
I then construct maps and faceted time series plots of the focal outcome variables, the nonemployment share and index crime rates, to describe the spatiotemporal variation in the outcome measures. The choropleth maps depict the mean level of each outcome variable averaged across the panel, and the maps are constructed using identical methods to that of the initial BTB status choropleth. The faceted time series plots display the temporal variation in each state’s nonemployment shares and crime rates across the panel, with each faceted plot corresponding to a different state. I color each state line based upon the current status and scope of the BTB legislation at each point in time.

Finally, I calculate pre- and post-treatment means for each scope of BTB law, to examine the change in each outcome variable pre-post a BTB law inception. I calculate means for overall nonemployment shares, both White and Black employment shares, as well as the total index crime rate. It should be stressed that these simple mean comparisons contain both variation that is ostensibly due to changes in treatment, as well as other factors correlated with time (e.g., declining overall crime rates). Therefore, a difference-in-difference research design will be used to isolate the variation that is due to the adoption of BTB legislation by comparing the temporal variation in treatment states to states that have not yet passed BTB legislation.

**Difference-in-Difference with Staggered Adoption**

To estimate the causal impact of BTB legislation on both nonemployment shares and crime rates, I use a difference-in-difference (DID) research design. The classic DID design comprises two time periods and two groups, where one group is treated and another group is not. The treated group becomes treated in the second time period and the difference in the change in the treated group as compared to the change in the untreated
group estimates the causal effect of treatment. Classic DID effectively estimates “how much more the treated group changed than the untreated group (Huntington-Klein 2021:437),” and the change in the untreated group represents the expected amount of change in the treated group had treatment been withheld, effectively acting as a counterfactual scenario.

Formally, the 2x2 classic DID estimates the average treatment effect among the treated (ATT) by comparing the observed, treated outcomes to the unobserved, but estimated outcome from observed untreated units:

$$ATT = E[Y_t(1) - Y_t(0) | D = 1]$$

where the ATT is the average of the difference between the treated, $Y_t(1)$, and untreated, $Y_t(0)$, potential outcomes for units in the treated group. The DID design relies upon the parallel trends assumption, which assumes that “had no treatment occurred, the gap between treated and untreated groups would have remained constant (Huntington-Klein 2022:441).” This assumption, if true, ensures that the difference between the observed treatment outcome and the unobserved, estimated counterfactual outcome can be attributed solely to the treatment, as the over-time changes in the absence of treatment would have been the same. The parallel trends assumption can be formalized as

$$E[Y_t(0) - Y_{t-1}(0) | D = 1] = E[Y_t(0) - Y_{t-1} | D = 0]$$

which effectively states that the change over time in the untreated counterfactual for the treated group, which is unobserved, is equal to change over time between in the untreated group, which is observed. This assumption allows the design to leverage the observed outcomes in the control as an estimate for the treatment units unobserved outcomes under
the condition of no treatment. It should be noted that the parallel trends assumption is often strengthened by conditioning on covariates, which can aid in controlling for differences over time that are not identical over time between treatment groups, making the parallel trends assumption more credible (e.g., Abadie 2005). Under the parallel trends assumption, the ATT is identified:

\[ ATT = E[Y_t - Y_{t-1}|D = 1] - E[Y_t - Y_{t-1}|D = 0] \]

which states that the difference in the mean change in outcomes over time in the treatment group is adjusted by the mean change in outcomes over time in the untreated group. Effectively, this “differences out” the part of the over-time change in the treated group that is not due to treatment.

As in many other applied settings, the classic 2x2 DID design does not fit the passage of BTB legislation laws, as these laws are adopted by different places in different times – what is called staggered adoption in the econometrics literature (Callaway and Sant’Anna 2021a). Staggered adoption designs deal with binary treatments (e.g., BTB law or no BTB law) that are adopted at different times across a panel, and once a unit participates in the treatment, the unit remains treated. This is the case for BTB laws, as no states to date have “rolled back” their BTB legislation after adoption as of the end of the study window (Avery and Lu 2021).

The most common estimation method for staggered adoption designs is the two-way fixed effects estimator (TWFE), which includes fixed effects for both units and time to control for the unobserved, time-constant heterogeneity between units, and the unit-constant, time heterogeneity unique to different treatment times. However, in contexts
with variation in treatment timing and heterogeneous treatment effects - either across time or across treated units-, the TWFE estimator makes contaminated comparisons between newly treated units to *already treated* units, effectively biasing the treatment effect with an injection of previous, possibly heterogeneous, treatment effects (e.g., Goodman-Bacon 2021, Sun and Abraham 2021). Put simply, fixed effect estimation compares *within groups*, so “still treated” units, which do not exhibit over time variation in treatment, get used as comparisons just as groups that are not treated do. This contaminates the comparisons between treated and untreated units, which is exacerbated when treatment effects are dynamic over time, a plausible scenario in the case of BTB laws (see Doleac and Hansen 2020).

Given these drawbacks of the TWFE model, and given the staggered rollout of BTB legislation in the U.S. coupled with the plausible treatment heterogeneity due to the differences between laws, I use a relatively new estimation technique called group-time average treatment effects (GTATT) (Callaway and Sant’Anna 2021a). GTATTs are essentially unique ATT estimations for a cohort of units treated at the same point in time. The GTATT estimation process is a two-step process which involves 1) estimation and inference about each of these disaggregated, unique cohort specific parameters, and 2) the aggregation of these parameters into summary measures of the effects of interest. This approach essentially computes individual “building block” 2x2 ATTs at each point in time for each treatment group (both pre- and post-treatment), which effectively make uncontaminated comparisons between treatment and control groups, making these estimates robust to treatment effect heterogeneity and dynamic treatment effects over time (Callaway and Sant’Anna 2021a). GTATTs can be expressed as
\[ \text{ATT}(g, t) = E[Y_t(g) - Y_t(0) | G = g] \]

which is the average effect of participating in the treatment for units in cohort treatment group \( g \) at time period \( t \). In the classic 2x2 DID setup, \( g = 2 \) and \( t = 2 \) gives the ATT. Extending the GTATT to period after treatment (e.g., \( \text{ATT}(2, 3) \)) estimates dynamic treatment effects, and extending a GTATT backwards in time (e.g., \( \text{ATT}(2,1) \)) calculates a “pseudo-ATT” that compares treated to untreated in the pre-treatment period, and can be used for the assessment of pre-trends to give evidence towards the validity of the parallel trends assumption (Callaway and Sant’Anna 2021a). Given the small size of the “never treated” group of states in the panel (\( n=16 \)), my staggered adoption design uses the “not yet treated” units as control units at each given time point. The parallel trends assumption for the design using the “not yet treated” units as controls, conditional on observed covariates, can be expressed as follows (Callaway and Sant’Anna 2021a):

\[ E[Y_t(0) - Y_{t-1}(0) | X, G = g] = E[Y_t(0) - Y_{t-1}(0) | X, G \neq g] \]

that is, that conditional on covariates \( X \), the change in outcomes for the treated units in the counterfactual untreated scenario is the same as the observed change over time in the untreated control units.

To estimate the GTTATs for BTB law passage, I utilize an outcome regression approach which regresses the outcome variable onto the included covariates, which creates a predicted value, using the data on untreated units, under the counterfactual scenario for the treated units (Heckman 1997; 1998). This value is then “differenced-out” of the observed, over time change in the treatment group to estimate the ATT for a given group at a given time. The model,
where \( D_{t+\delta} \) is an indicator whether the unit has as of yet received treatment, \( \delta \) is the anticipation window, \( X \) is the covariates to adjust for, and \( Y_t - Y_{g,t,\delta} \) represent the difference between the outcome at time \( t \) and time \( g - \delta - 1 \) (i.e., the baseline comparison period). I utilize an anticipation window of \( 1 \) across all models, to remove any potential 1-year anticipation effects that may appear due to states “choosing” to pass BTB legislation. The ATT is then estimated with the following equation (Callaway and Sant’Anna 2021a):

\[
ATT_{or}(g, t; \delta) = E \left[ \frac{G_g}{E[G_g]} (Y_t - Y_{g,t,\delta} - 1 - Y_{g,t,\delta}(X)) \right]
\]

which estimates the average ATT across units in group \( g \) in time \( t \). Inference on the GTATTs, as well as aggregations of the GTATTs, are performed via a multiplier bootstrap procedure which results in a simultaneous confidence band that is robust to multiple comparisons (see Callaway Sant’Anna 2021a:24 for details).

The estimation and inference of each GTATT results in the estimation of a 2x2 ATT (or pseudo-ATT if pre-treatment), for each treatment group for each year. This results in the estimation of a large number of estimated parameters. For example, the first specification of first BTB law passage of any scope results in the estimation of 242 individual GTATTs. Callaway and Sant’Anna (2021a) detail methods for summarizing the parameters via a variety of methods to examine overall average effects, as well as treatment effect heterogeneity and dynamic treatment effects over time. The first is what I call the “overall” aggregation technique, which can be expressed as follows:
\[ \theta_{overall} = \sum_{g \in G} \frac{1}{T-g+1} \sum_{t=g}^{T} ATT(g, t) P(G = g | G \leq T) \]

where GTATTs are first summed over treatment groups across all time periods, and then the group-specific effects are averaged together, weighted by the number of units in each group, across groups to arrive at a single estimate of BTB law adoption. This overall aggregation estimates “the average effect of participating in the treatment experienced by all units that ever participated in the treatment (Callaway and Sant’Anna 2021a),” and therefore can be interpreted akin to the ATT in a classic 2x2 DID design. By first summing over the group effects, it avoids the problem of weighting groups that experienced treatment earlier more heavily, as would happen in a simple summation of post-treatment ATTs (see Callaway and Sant’Anna 2021a:18).

The second aggregation method utilized in the analysis here mimics event study regressions, which include interactions between the treatment variable and time leads and lags to examine treatment effect dynamic both pre- and post-treatment. The “event study” aggregation method, is expressed as follows:

\[ \theta_{es(e,e')} = \sum_{g \in G} 1\{g + e' \leq T\} \sum_{t=g}^{T} ATT(g, g + e) P(G = g | G + e' \leq T) \]

where the ATTs for time period \( g + e \) is calculated for each treatment group, and then aggregated across groups in relative event time. \( e' \) is equal to \( e \), and denotes units who are observed to have participated in treatment for at least \( e' \) periods, which restricts the calculations of dynamic treatment effects to just the units that have experienced treatment of the specified lag. This effectively balances the comparisons to include the same
treatment units across each lead/lag aggregation, which avoids the issue of compositional changes across relative event time at the cost of fewer groups being used to estimate the event study parameters (Callaway and Sant’Anna 2021a). These aggregations are useful for examining dynamic treatment effects across relative event time, as well as determining if pre-treatment trends, which would call into question the parallel trends assumptions, exist. Given the predominance of BTB law passage in the most recent decade in the study period (2010s), I calculate aggregated event study parameters for leads and lags of three years pre- and post-treatment, and balance the comparisons to 3 observed post-treatment periods.

The final aggregation method I utilize is a simplification of the “overall” method, which aggregates the ATTs by treatment group:

$$\theta_{group} = \frac{1}{T - g + 1} \sum_{t=g}^{T} ATT(g, t)$$

which effectively averages the ATTs for the units that experienced treatment in the same year across all post-treatment periods. These parameters are useful in examining treatment effect heterogeneity between treatment groups.

In each model, I include a bevy of covariates which could have contemporaneously changed with treatment rollouts, and have been shown to be key drivers of nonemployment (see Larson et al. 2021), as well as pattern selection into BTB policy adoption (Doleac and Hansen 2017). The inclusion of the covariates adjusts the counterfactual scenario for between treatment group differences on these measures, making the conditional parallel trends assumption more credible. The notes below each
regression table document the covariates adjusted for in the estimation of each group of GTATTs, as the covariates included vary slightly across specifications.

I first estimate a nonemployment specification defining treatment as the first passage of any BTB legislation, regardless of the scope of the law. This serves as the primary specification, and I subsequently estimate nonemployment specifications that define treatment differently, in order to examine the heterogeneity of different BTB laws in terms of scope. The second nonemployment model defines treatment as a full public ban or better, leaving states that had were untreated, or only passed state employment bans, as the untreated group. The third specification defines treatment as a full public and private combination, consisting of the treatment date of states that either a) passed a full public and private ban in tandem, or b) passed full private ban after already having a public ban in place.

I then estimate separate specifications for White and Black nonemployment, to examine the extent to which BTB laws may be differentially efficacious based upon the race group under consideration. I also estimate separate DIDs with split samples of states consisting of those above and below the mean level of felony history as of 2010, to evaluate whether the share of record holders not currently under correctional supervision moderates the impact BTB law passage has on nonemployment. Finally, I estimate a crime rate GTATT specification, examining the impact of BTB legislation on state-level crime rates. All models are estimated with the ‘did’ package (Callaway and Sant’Anna 2021b) in R. It should be noted that in interpreting the DID results below, I focus primarily on effect magnitude, as opposed to statistical significance or confidence.
intervals, as the models are estimated based upon a population of U.S. states in the years 1995-2020.

Results

Descriptive and Bivariate Results

The first part of the analysis of this empirical chapter concerns the spatial distribution, and over time passage, of BTB legislation in the United States from 1995-2020. Figure 12 displays the distribution of BTB policies as of 2020, the end of the study period. As of 2020, 15 states had still to pass any form of BTB legislation, including states such as Texas, Alabama, and Florida. This represents the modal category of BTB status in the U.S., as the majority of states have started to pass BTB in the mid to late 2010s (see Table 12 and Figure 13). In contrast, 14 states have passed a state employment ban, without passing any subsequent broadening of the state’s law. This represents second most frequent category of BTB law status, as these states have not expanded the ban past state level employment. Nine out of 50 states remain with a full public ban covering both state and lower-level municipal employment (e.g., Ohio, Virginia, Nevada), and a further 12 have moved to a full public and private BTB regime as of 2020 (e.g., Minnesota, California, Colorado). Illinois serves as the lone outlier, having passed a state-level BTB ban in 2013 followed by a full private ban in 2015.26 In sum, BTB legislation has become a normative policy for states to pass, particularly in the latter part of the study period (see Figure 13), but legislation still varies significantly across states in terms of the scope of the employment the bans apply to.

26 Although outside the study window, Illinois passed a full public and private BTB legislation becoming “just the second state (after New York) to extend antidiscrimination rights to workers with conviction records (Avery and Lu 2021:13).”
Consistent with Figure 12, the Kaplan-Meier survival plot in Figure 13 displays the proportion of states existing in the “control” state for each level of BTB treatment indicator. The red line is the survival curve for any BTB law passage. As of the end of the study period, 30% of the states had “survived” without passing any form of BTB legislation, with the majority of states beginning to pass legislation around 2010, with sharper increases in failure – or BTB adoption – around 2015. The survival curve for at least a full public ban follows a similar temporal trajectory, but fewer states have reached this treatment threshold, with over 60% of states having yet to pass a ban beyond state employment. Finally, the most restrictive ban, a full public and private ban, exhibits the greatest proportion of survival at the end of the study period, with over 75% of U.S. states not reaching legislation that applies in any form to private employment. These survival estimates show that less restrictive laws were more likely to be passed and more widely adopted, and the timing of BTB legislation was accelerated towards the latter half of the study period.
Next, I turn to describing the spatial and temporal variation in the focal outcome variables: prime-age nonemployment shares and the total index crime rates per 100,000. First, Figure 14 maps each state’s prime-age nonemployment share as of the end of the study period, 2020. In general, the upper Midwest exhibits lower levels of nonemployment as compared to other parts of the U.S., with a few states, such as Nevada, Hawaii, and Mississippi standing out as having particularly high nonemployment shares as of the end of the study period (> 29%).
Figure 15 displays a choropleth map of the UCR index crime rate per 100,000, the second focal outcome variable in this analysis. At the state-level, the states in the Northeast region display slightly lower index crime rates per 100,000. In contrast, nine states (e.g., Louisiana, New Mexico, Colorado) report index crime rates above 3000 index crimes per 100,000 population aged 16 or over.
Figures 16 and 17 individually plot each state's outcome trends over time, with lines segments colored by the current BTB law scope in place at the time. This allows a visual evaluation of the post-treatment trends in each outcome variable as compared to periods of post-treatment. In terms of the nonemployment share, states generally follow similar trends with an increase in the middle of the study window (following the stock market crash in 2008), with general decreases until 2020, where the COVID-19 pandemic led to increased levels of unemployment (Auginbaugh and Rothstein 2022).

Looking visually at the potential treatment effects, the majority of nonemployment shares show declines in years post-treatment, which is consistent with the hypothesis that BTB legislation bolsters employment. However, states also tended to adopt BTB legislation post-2010, a period marked by declining nonemployment. Similar to the trends in the nonemployment shares, the index crime rates show general decreases post-treatment, which is evidence supportive of BTB legislation reducing crime rates. However, crime rates, across the vast majority of states, have been gradually declining since the start of the study period, making treatment colinear with general decreases in state-level crime. While these descriptive time series plots provide a visual of the outcome trends, as well as the trends after treatment adoption, a more rigorous evaluation is whether states that adopted BTB legislation had more pronounced decreases in the outcomes as compared to similarly situated untreated states, which the staggered adoption DID models below evaluate.
I also examine the bivariate association between BTB law adoption with mean comparisons between state-years in post-treatment statuses, as compared to state-years in pre-treatment, or never treated, conditions (Table 14). In terms of state BTB bans, the mean nonemployment shares are slightly higher post-treatment across overall, White, and Black nonemployment shares. Similarly, full public bans or better exhibit slightly higher overall, White, and Black nonemployment shares post-legislation. In contrast, the full
public and private combination bans show lower nonemployment shares (across all three types) post-treatment, which may suggest that more restrictive BTB legislation that includes private employment may have greater impacts on macro-level employment. Finally, all state-years in post-treatment conditions exhibit significantly lower index crime rates as compared to state-years pre-treatment or never treated. It should be stressed that these simple pre/post mean comparisons contain both variation that is could be due to changes in treatment, as well as other factors correlated with time (e.g., declining overall crime rates). The focal one here being the onset of the COVID-19 pandemic in early 2020, which is colinear in time with post-treatment for the treated state-years. Therefore, the higher post-treatment means could be, at least in part, due to the significant increase in nonemployment at the end of the study window due to the start of the COVID-19 pandemic. Therefore, a difference-in-difference research design will be used to isolate the variation that is due to the adoption of BTB legislation by comparing the temporal variation in treatment states to states that have not yet passed BTB legislation, which removes the time-stable impact of COVID-19 from the treatment effect.

<table>
<thead>
<tr>
<th>Legal Scope</th>
<th>Pre/Post Legislation</th>
<th>Prime Nonemployment Share</th>
<th>Prime Nonemployment Share - White</th>
<th>Prime Nonemployment Share - Black</th>
<th>Total Index Crime Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>Pre</td>
<td>20.18</td>
<td>20.18</td>
<td>26.66</td>
<td>3370.82</td>
</tr>
<tr>
<td>State</td>
<td>Post</td>
<td>21.69</td>
<td>21.69</td>
<td>27.48</td>
<td>2163.14</td>
</tr>
<tr>
<td>Public-</td>
<td>Pre</td>
<td>19.16</td>
<td>19.16</td>
<td>25.03</td>
<td>3599.88</td>
</tr>
<tr>
<td>Public-</td>
<td>Post</td>
<td>20.84</td>
<td>20.84</td>
<td>26.71</td>
<td>3161.69</td>
</tr>
<tr>
<td>Public/Private</td>
<td>Pre</td>
<td>20.31</td>
<td>20.31</td>
<td>26.16</td>
<td>3346.64</td>
</tr>
<tr>
<td>Public/Private</td>
<td>Post</td>
<td>19.89</td>
<td>19.89</td>
<td>26.07</td>
<td>2677.37</td>
</tr>
</tbody>
</table>

**Staggered Adoption Difference-In-Differences**

I now turn to present the staggered adoption DID models, which estimate the group-time ATTs, and I aggregate them following the methods outlined in the Analytical
Strategy section. Figure 18 displays a table of the ATT aggregations of the first treatment threshold, a state’s first BTB law of scope, as well as visualization of the event study and group aggregations. The overall ATT estimate is -1.69, indicating that the passage of any BTB legislation causally reduces the overall nonemployment share by ~1.7 percentage points. The event study pre-treatment pseudo-ATTs hover around 0, indicative of no pre-trends before treatment (Callaway and Sant’Anna 2021a). Post-treatment, the event-study lags exhibit a growing negative effect in the first three years post-treatment followed by an attenuation towards zero in year four. In the group aggregations, which displays the treatment effect heterogeneity across different treatment groups, all except one are negative, ranging in magnitude from -6.61 to -.22 percentage point declines in the periods post-treatment. This primary specification indicates that the average causal effect of BTB legislation is negative, with some variation in how effective various BTB laws can be. I now turn to delve into this treatment effect heterogeneity more deeply by considering alternative treatment thresholds.
Figure 18: Staggered Adoption DID Group-Time ATTs of First BTB Law

<table>
<thead>
<tr>
<th>Age Type</th>
<th>Lead Lag</th>
<th>Group</th>
<th>ATT</th>
<th>SE(ATT)</th>
<th>LCL</th>
<th>UCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>--</td>
<td>--</td>
<td>-1.6911262</td>
<td>0.6033609</td>
<td>-2.8697817</td>
<td>-0.5124708</td>
</tr>
<tr>
<td>Event Study</td>
<td>-3</td>
<td>--</td>
<td>-0.0327927</td>
<td>0.2746025</td>
<td>-0.5710036</td>
<td>0.5044182</td>
</tr>
<tr>
<td>Event Study</td>
<td>-2</td>
<td>--</td>
<td>-0.1224331</td>
<td>0.3004368</td>
<td>-0.7120886</td>
<td>0.4650200</td>
</tr>
<tr>
<td>Event Study</td>
<td>-1</td>
<td>--</td>
<td>0.2183704</td>
<td>0.3536600</td>
<td>-0.4685662</td>
<td>0.9323306</td>
</tr>
<tr>
<td>Event Study</td>
<td>0</td>
<td>--</td>
<td>-0.3127063</td>
<td>0.4817061</td>
<td>-1.0983290</td>
<td>0.7942204</td>
</tr>
<tr>
<td>Event Study</td>
<td>1</td>
<td>--</td>
<td>-0.2297349</td>
<td>0.4348240</td>
<td>-1.0819763</td>
<td>0.6225026</td>
</tr>
<tr>
<td>Event Study</td>
<td>2</td>
<td>--</td>
<td>-0.7215857</td>
<td>0.4331272</td>
<td>-1.5705024</td>
<td>0.1273253</td>
</tr>
<tr>
<td>Event Study</td>
<td>3</td>
<td>--</td>
<td>-0.3316122</td>
<td>0.7252525</td>
<td>-1.6682706</td>
<td>1.3499661</td>
</tr>
<tr>
<td>Group</td>
<td>--</td>
<td>2009</td>
<td>-0.0486878</td>
<td>0.5562433</td>
<td>-5.1591005</td>
<td>-2.9187746</td>
</tr>
<tr>
<td>Group</td>
<td>--</td>
<td>2010</td>
<td>0.7050706</td>
<td>0.7830784</td>
<td>-0.8297348</td>
<td>2.2398761</td>
</tr>
<tr>
<td>Group</td>
<td>--</td>
<td>2013</td>
<td>-0.6029709</td>
<td>0.6450559</td>
<td>-5.7802662</td>
<td>-5.346936</td>
</tr>
<tr>
<td>Group</td>
<td>--</td>
<td>2014</td>
<td>-0.2519523</td>
<td>1.6472739</td>
<td>-3.4437899</td>
<td>3.0134053</td>
</tr>
<tr>
<td>Group</td>
<td>--</td>
<td>2015</td>
<td>-0.9783934</td>
<td>1.0496916</td>
<td>-3.0362608</td>
<td>1.0794739</td>
</tr>
<tr>
<td>Group</td>
<td>--</td>
<td>2016</td>
<td>-0.7178991</td>
<td>0.5823357</td>
<td>-1.8159250</td>
<td>0.4234879</td>
</tr>
<tr>
<td>Group</td>
<td>--</td>
<td>2017</td>
<td>-4.4173737</td>
<td>2.7519219</td>
<td>-9.8172414</td>
<td>0.9817701</td>
</tr>
<tr>
<td>Group</td>
<td>--</td>
<td>2018</td>
<td>-1.4425644</td>
<td>1.0997333</td>
<td>-3.5980020</td>
<td>0.7197332</td>
</tr>
</tbody>
</table>

Note: Anticipation = 1, controls are unit not yet treated. DID specification includes covariates for population age shares, marriage, bachelor’s degree, disability, and poverty rates, the effective minimum wage, AFDC/TANF/SNAP Max, the unemployment rate, and three year lags of the unemployment rate.

Figure 19 displays the staggered adoption DID model results from a specification that defines treatment as a full public ban or more restrictive (Public +), as compared to no ban or only a state level ban. Therefore, the treatment effect for a state-level ban is rolled into the “control” state, and therefore the ATT here would represent the impact of a public or better ban above and beyond any state employment treatment effect. The overall aggregated ATT exhibits a similar causal effect to the primary specification, with a public
ban or more restrictive leading to a reduction in the overall nonemployment share of 1.29 percentage points. This suggests that the evolution to a more restrictive ban can garner further returns to employment on top of a state BTB law. The event study pre-treatment aggregations do not show a pattern in pre-trends before treatment although the 1-year lead does move in the direction of treatment, which could indicative of anticipation of treatment longer than 1 year. The post-treatment lag aggregations show similar patterns to the primary specification, with initial decreases in nonemployment in the three years immediately following treatment, with a subsequent reduction in efficacy in year four. Additionally, the estimated group ATT aggregations are primarily negative, suggesting that full public bans, or bans that reach both public and private employment, aid in reducing nonemployment.
Figure 20 relays the group-time ATT aggregations for the treatment threshold of a full public and private ban, which defines treatment as the passage of a full public and private BTB law, folding less restrictive bans into the “control” group. In contrast to the primary and public-plus treatment thresholds, the ATT for the most restrictive BTB law is negative, but much smaller than that of the other specifications, with treatment causing a -.22 percentage point reduction in nonemployment. This suggests that the addition full
public bans do not significantly reduce nonemployment above and beyond the control group (which includes states with full state or public bans), and is consistent with some previous work that finds weak effects of additional public bans (Doleac and Hansen 2020). The event study aggregations pre-trends indicate that the pseudo-ATTs do not exhibit a pattern pre-treatment. The post-treatment event study aggregated ATTs show a similar pattern to the other specifications with initial decreases in nonemployment followed by a return towards zero. In contrast to the previous specifications, the group aggregations ebb and flow around zero, giving further evidence that this treatment threshold exhibits weak effects on nonemployment. In sum, the treatment threshold specifications indicate that BTB legislation has the ability to improve employment outcomes for states, and that full public bans may be particularly effective. However, the third treatment threshold of full public and private bans suggest that adding a private ban does not result in significant returns to employment. I return to why this may be in the discussion section below.
I now turn to the conditional primary specifications (Table 15), which will examine the treatment heterogeneity amongst different racial groups, as well as states with different levels of felony criminal record production. The White nonemployment specification reveals a -.32 causal effect of any BTB legislation. In contrast, and in concert with previous scholarship (e.g., Doleac and Hansen 2020), the Black nonemployment specification exhibits a much larger and positive causal effect, with over a 7-percentage point increase in nonemployment amongst Black individuals. This ATT is
very large in magnitude, and caution is warranted, as the standard error of the ATT is relatively high. This instability may be present due to the small treatment groups and the more variable Black nonemployment rate. This instability could also be due to missing data amongst some state-years in the CPS basic monthly data, as the estimation is performed on a balanced panel. With this caveat, this finding may suggest that, on average, employers, in the presence of more limited criminal history information, discriminate on the basis of race to higher degree due to stereotypes between Blackness and criminality (e.g., Quillian and Pager 2001) and I return to this in more detail in the discussion to follow.

The split population felony history specifications show that the efficacy of BTB legislation is also contingent upon a state’s history of felony criminal record production. BTB legislation exhibits a positive effect in states with above mean felony history percentages (4.04), and a negative, and lower in magnitude, effect in states with below mean historical felony record production (-.78). This may suggest that employers, dealing with uncertain information in the employment process, may use proximal information on the prevalence of criminal records, akin to the race effects above, to “hedge their bets” in terms of criminal history. Alternatively, this could reflect that the laws are not sufficient to create a dent in the unemployment status of those with felony level criminal records in states with high prevalence, as compared to states with lower levels of extant felony criminal record holders. Again, I return to this in greater detail in the discussion section to follow.
The final staggered adoption DID specification in Table 21 examines the causal
effect of BTB legislation on state-level index crime rates. Consistent with the primary
nonemployment specification above, BTB law adoption exhibits a negative effect on aggregate crime rates, with adoption leading to a drop in 68 index crimes per 100,000. Given the fairly large standard deviation of the total index crime rate of 1158.4, this represents an effect of relatively small magnitude as compared to the effect of BTB on nonemployment. This makes some intuitive sense: if employment is the primary pathway by which BTB legislation impacts crime, I should expect the effect on crime to be lower as the effect of employment on crime at the state-level is relatively weak (Levitt 2001; Arvanites and DeFina 2006) and sand even weaker in terms of violent crime (Raphael and Winter-Ebmer 2001). Despite being a different outcome, the event study ATT aggregation show strikingly similar patterns to that of the nonemployment specifications, with initial post-treatment decreases, albeit modest in magnitude, attenuating toward 0 by year 4. Additionally, the group ATT aggregations are roughly split between positive and negative ATTs, suggesting that BTB legislation may not reduce state-level crime rates. Overall, the results of the crime rate DID specification provide weak, at best, evidence for a causal relationship with BTB legislation, although the overall aggregated ATT is in a crime reducing direction.

**Discussion and Conclusion**

This empirical dissertation chapter aimed to examine the causal effects of BTB legislation on both employment and crime, thereby evaluating a state policy with the theoretical potential to “cut off” part of the iatrogenic reach of punishment on crime. This research not only has important policy implications for both employment and crime, but also illustrates sociologically how policies can have heterogeneous effects that may be counterintuitive.
First, this chapter corroborates previous DID studies that BTB legislation can have ameliorate effects on overall employment. Specifically, the primary any law specification yielded a causal estimate of -1.7 percentage points, which is fairly substantial considering that the laws only target a minority, albeit growing, subset of the American population (Shannon et al. 2017). Moreover, this chapter contributes to the literature by explicitly parameterizing different treatment thresholds. While I find an effect of significant magnitude for states that pass a public ban or more restrictive (ATT = -1.7), the full public and private combination treatment threshold resulted in an effect about a tenth the size of the previous specifications. This suggests that the efficacy of BTB legislation in terms of moving the needle of employment happens in the move to full public bans, with minimal returns given to moving towards placing restriction on private employment. This result is surprising, especially considering that one would expect that the BTB legislation with the highest “treatment dosage” would further enhance employment. However, previous scholarship has documented the lack of compliance amongst employers, estimating lack of compliance with BTB laws at a rate around 20%, with noncompliance more prevalent amongst employers who discriminated against those with criminal record pre-BTB (Schneider et al. 2021). While this study does not examine variation in public vs. private employers in terms of compliance, it could be the case that private employers are less compliant with BTB legislation, and therefore the potential additional gains of the private bans on top of other forms of BTB legislation are weak. Although speculative, future inquiry, with the right data, could tease out whether this is an explanation for the lack of nonemployment reduction in the full public-private model.
This chapter also corroborates previous research that finds heterogeneous treatment effects of BTB by race. While finding a significant negative effect on overall employment, the race-subgroup DID specifications indicate a negative effect for White nonemployment, but a positive effect on Black nonemployment. This is consistent with previous scholarship (Agan and Starr 2018; Doleac and Hansen 2020), as well as statistical discrimination arguments wherein employers leverage their preexisting stereotypes of race and criminality in the absence of criminal history information when making employment decisions. However, the magnitude of this effect and its large standard error suggests that caution is warranted in its interpretation. Some instability in the effect may be present due to the small treatment groups and the comparably more variable Black nonemployment rate.

With this caveat, these results represent challenges to BTB legislation, as growing evidence has marked BTB as having harmful unintentional effects on communities already subject to intense forms of structural racism inherent in the criminal legal system (e.g., Alexander 2012). Vuolo et al. (2014) discuss the delicate balance that BTB laws need to strike to reduce racial discrimination in employment, as to minimize the harmful effects on those without records while simultaneously reducing the discrimination against those with records. Vuolo et al. (2014) suggest that BTB legislation that prevents criminal record checks until the interview or offer stage may catalyze the effects of “getting a foot in the door (159),” which may help offset practices of statistical discrimination on the basis of race. Sugie (2017) recommends that instead of “wavering” on BTB policies due to statistical discrimination concerns, strategies should be adopted to contest race-based discrimination in the labor market. In addition, two states, New York
and Illinois, have extended antidiscrimination rights to workers with criminal records. While too soon to examine the impact of bringing criminal records into the realm of civil rights, more aggressive laws such as these may alter the statistical discrimination that takes place during the hiring process. Thus, it remains an open question as to what BTB laws are least amenable to statistical discrimination.

Another counterintuitive finding concerns the felony history DID specifications, which show that BTB legislation is far more efficacious in reducing nonemployment in states with lower felony history shares. These findings run counter to what has been found previously, with more efficacious treatment effect found amongst places with a higher prevalence of criminal records (e.g., Shoag and Veuger 2021). It is possible that state-level felony history works in a similar fashion to that of race: places with a high proportion of criminal records may exhibit lower effects in response to BTB legislation, as criminal records are more prevalent in the local labor market and employers account for this proximal information in the hiring process. Put differently, employers may be structurally given more latitude to exclude individuals with criminal records when these applicants are plentiful relative to the number of available positions. Another possibility is that BTB laws are not a strong enough treatment to combat the prevalence of criminal records in some high record production states, as employers can still choose to discriminate on the basis of a criminal record after a certain stage of the employment process in most states (Raphael 2021).

Finally, this chapter makes a contribution to both the BTB and punishment effects literatures by considering crime as a focal outcome of BTB legislation. Unlike the overall nonemployment specification, the results of the crime DID were weak, albeit in the
expected negative direction. As stated in the literature review, this is consistent with previous work that shows that the effect of employment on crime at the state-level is relatively weak (Levitt 2001; Arvanites and DeFina 2006) and has relatively weaker relationships with violent crime as compared to property crime (Raphael and Winter-Ebmer 2001). Therefore, if the primary “lever” by which BTB can impact crime is through employment, the distal outcome of crime is not as responsive to policy adoption. Given the myriad of factors that comprise the etiology of crime, it is not all that surprising that a policy that selectively targets only one of the potential causal pathways between punishment and crime show relatively minimal impact. In sum, BTB laws as currently practiced do not have the potency to massively alter employment, and therefore have an even lesser impact on crime given the relatively weak, albeit negative, relationship between employment and crime at the macro level.

Although this chapter conducts rigorous statistical analyses of the effects of BTB legislation, the analyses are not without their limitations. The results here are estimated solely on state-level ban the box legislation, and therefore do not apply to BTB legislation at more local jurisdictions (e.g., cities). The effects here, although estimating the causal effect of state-wide bans, are also potential underestimates of the full reach of BTB legislation, as lower-level BTB ordinances may contribute to state-level employment and crime above and beyond state-level bans. Additionally, the time limit on the felony history share data could lead to misleading conclusions if the relative rank of states in terms of their felony history share has changed drastically over time. The measurement of crime is always a tricky endeavor, and the UCR data used here are subject to potential reporting biases that induce measurement error into the crime rates
(Kaplan 2021a), not even considering crimes that are not reported to the police. These problems are partially mitigated by the use of more serious index crimes, which are measured more accurately (apart from rape, Kaplan 2021a), than less serious forms of crime. Finally, while the event study ATT aggregations were largely free of noticeable pre-trends before treatment, the possibility always exists in DID designs that the parallel trends assumption breaks down after treatment, which is something that is not testable, as the treatment counterfactual is unobserved.

Despite the caveats listed above, this empirical dissertation chapter provides rigorous quantitative evidence of the impacts of BTB laws, with significant impacts on employment. It is also the first study to examine BTBs impact on crime, finding weak BTB treatment effects on crime. This suggests that while BTB may have some ameliorative capacity in terms of employment, the evidence does not suggest that BTB laws are strong enough to reduce the iatrogenic effects of records on crime. Further, it has detailed the heterogeneity in the effect of BTB laws by scope, with full public bans or better generally having the most pronounced effect on employment. However, this chapter also cautions that BTB can have harmful effects on Black employment, and I conclude that BTB laws are not a perfect solution to the problem of criminal records in the labor market. Broadly, this shows that, while well intentioned in its manifest goals, BTB legislation may have unintended latent consequences that are detrimental and diligence should be fostered when implementing these policies. While limited in their impact and scope in terms of improving public safety, BTB laws show some promise for reducing the harmful effects of records in the labor market, and have the potential to improve the economic and social lives of those with criminal records. While BTB
legislation does reduce aggregate employment, policies that work to reduce criminal
record production, or limit access to criminal records altogether (Sandoval and Lageson
2022; Lageson 2020), may provide even more efficacious avenues for reducing the
deleterious effects of records in the labor market. Finally, while BTB laws have the
potential to reduce racial disparities in employment and other economic outcomes by
removing barriers to employment, a countervailing effect of BTB may work to
undermine these positive gains for Black individuals. Future research should examine
potentialities of how to improve the former, whilst reducing or eliminating the effects of
the latter.
Conclusion: Sociological Insights, Policy Implications, and Future Directions

This dissertation has detailed three empirical studies that examine various aspects of the iatrogenic link between crime and punishment across multiple levels of analysis. While no one study, or dissertation, cannot adequately describe the entirety of the punishment-crime nexus, these studies give insight into how punishment can impact the distal outcome of crime in complex and often counterintuitive ways, as well as highlight the paradoxical nature of current criminal legal practice in its attempts to control crime. Further, this dissertation speaks to the efficacy of state policy in alleviating some of the adverse effects punishment can have on social life. Each of the empirical studies within this dissertation contain insights for future sociological research in the realm of crime and punishment, as well as policy implications for punishment in American society.

The studies within this dissertation are situated in a long line of sociological inquiry that highlights both the manifest and latent functions (Merton 1957) of social processes and institutions. The first chapter illustrates how punishment’s effects at their most direct are rather weak, and that colloquially “less serious” forms of punishment, often neglected in both concern and study to custodial forms of punishment, can have pernicious iatrogenic effects. Further, the results suggest that punishment should be analyzed more closely to how it is experienced. I regard this as a fundamental sociological insight, in that the package of punishment can have impacts above and beyond that of any punishment in isolation. Given that punishments tend to be experienced together, these findings suggest that criminological inquiry should analyze punishment more closely to how it is experienced by defendants.
The second empirical chapter examines how the spatial concentration of punishment into neighborhoods adversely impacts community crime by worsening disadvantage in communities. This also gives a sociological expansion to theories of neighborhood ecology, with punishment having the capacity to shape neighborhood characteristics above and beyond other facets of neighborhood structure. The findings highlight the bifurcating effects punishment can have at the community-level, having both direct and indirect effects that counteract one another in shaping crime. This research further implicates punishment as a prominent facet of structural racism alongside (and related to) racial segregation (e.g., Krivo et al. 2009) in American communities. The patterns described in this dissertation chapter illustrate punishment’s capacity to reproduce, and perhaps exacerbate, levels of community disadvantage and crime, and therefore characterizes punishment as operating alongside other forces within structural racism that (re)produce concentrations of race and space. While the study only documents one iatrogenic pathway by which punishment translates to crime at the community-level, questions remain as to the other perverse implications concentrations of punishment may hold for communities that also work in a similar countervailing fashion in terms of crime.

The BTB study is also within the sociological tradition of identifying the diverse effects of institutional policies. The chapter suggests that while BTB legislation may be able to ameliorate some of the barriers to employment amongst those with records, it may actually hinder the employment prospects for those already most vulnerable in the labor market (Pager, Bonikowski, and Western 2009). This highlights that evaluations of policies aimed at the iatrogenic effects of punishment must be attentive to whom the
policy helps, and those for whom the policy may not be effective or exacerbate the problem at hand. Further, the study is the first to show that BTB legislation may not be potent enough to completely sever the iatrogenic links between criminal records, employment, and crime, as the policies were shown to be less efficacious amongst states with higher prevalence of criminal record holders, and state-level crime rates were not responsive to BTB adoption. Given that employers can still discriminate against criminal record holders even in the presence of a BTB policy in the vast majority of states (Raphael 2012), BTB legislation may be a relatively weak treatment that lacks the necessary reach to make an observable dent in aggregate crime. Overall, each of these studies shares in the sociological tradition of documenting the varied, and sometimes counterbalancing impacts of social and institutional processes.

These studies also hold policy implications for both punishment practice as well as legislation designed to alleviate the strains of punishment on social life. The packaging effects revealed here implicate the tendency of the criminal legal apparatus to “pile on” (Uggen and Stewart 2015) those processed in the criminal legal system as a problematic criminogenic practice. While only looking at three focal aspects of punishment, the study calls into question current punitive practice of combining various punishments, requirements, and restrictions together as worsening not only crime but potentially other outcomes as well. Further, the first empirical chapter is suggestive that policies that “equalize” the courtroom playing field in terms of defendant disadvantage could be a possible route by which punishment’s adverse effects can be lessened.

The second empirical chapter examines how the current punishment practice of spatially concentrating forms of punishment into certain communities affects community
well-being and crime. The individuals punished within these communities often serve multiple social roles (e.g., parents, teachers), and “removing these residents eliminates their actual and potential role in neighborhood self-regulation (Rose and Clear 1998:469).” Thus, punishment can, on the one hand, perform a public safety function via incapacitation, but can also destabilize communities, dimmish social capital, and exacerbate local levels of disadvantage and crime. These patterns suggest that formal institutions of punishment must not only reckon with the damage that their practice causes in terms of community disadvantage, but also work to alleviate the harms caused by socially and spatially concentrating the burden of punishment. Policies that provide mechanisms for bolstering neighborhood social cohesion and collective efficacy, such as neighborhood organizations (Sharkey et al. 2017), as well and non-criminal-justice approaches to crime, such as restorative justice approaches (Van Ness and Strong 2014), may help strengthen the ability of communities to deal with the harmful effects of criminal victimization while simultaneously avoiding the social damage wrought by punitive responses to crime. Such findings are consistent with the classic Chicago School emphasis on social organization and investment in community resources to enhance neighborhood social cohesion.

The third paper arguably imparts the most direct prescriptions for state policy. The study indicates that while BTB laws may have the potential to improve the labor market prospects amongst those with records overall, it may work to undermine these positive overall gains by worsening employment prospects for Black Americans. Thus, public policy that attempts to ameliorate the barriers criminal punishment and its resulting stigma of a criminal record have on social life must be attentive to not only the
magnitude of its impact, but also for whom these policies work. Further, the lack of effect found in the crime rate specifications make sense intuitively, as BTB legislation only targets one of the many collateral consequences of punishment. This suggests that ex post facto policies that are limited in scope in terms of ameliorating punishment’s social, political, and economic consequences may not be effective in combatting the iatrogenic reach of punishment. These findings suggest that reducing the application of criminal records may have beneficial effects across various domains of social life (e.g., employment), and are consistent with the logic of diversion approaches to criminal justice (Wright and Levine 2021). Beyond altering the levels of punishment and production of criminal records, another potential fruitful alternative route to alleviate the harmful effects of records in the labor market is the wholesale reimagining of how criminal history information is gathered, disseminated, publicized, and used in decision-making (Sandoval and Lageson 2022; Lageson 2020). Rather than trying to “correct” for the stigmatizing effects this information can have with policies such as BTB, restricting the access and use of this information (i.e., akin to how medical information is handled) may be a more efficacious strategy to eradicate some of the social harms of punishment.

The research here also illustrates a path forward for sociological criminology as a field. While it has proven fruitful to identify the vast array of impacts of punishment on social life, the implications these relationships hold for crime should be documented as well. This helps reveal the ways in which punishment may not only be adversely impacting the lives of individuals within its grasp, but also counterproductively harming public safety in the process - potentially furthering justification for the current severity and scale of punishment. Given the consequences of criminal victimization for the social
well-being of communities (e.g., Sharkey 2018a), sociological and criminological inquiry would do well to take stock of the “social ledger of punishment” in regards to crime. This involves describing the multitude of social relationships within the nexus of punishment and crime and also trying to gain purchase on calculating more accurately and fully the balance of punishment’s import for crime. Given the racialized character of modern American punishment (e.g., Alexander 2012), this inquiry also has the potential to detail how the criminogenic harms of punishment are unequally experienced along the lines of race (e.g., Peterson and Krivo 2010), and reveal iatrogenic causal pathways that may feedback into popular public opinion that buttress the political fervor not only for punishment as a whole, but punitiveness expressed towards racial minority groups (Enns 2016; Chiricos et al. 2004). In summation, the iatrogenic effects under study in this dissertation are archetypal of research that can unite the inquiry of the social harms of punishment with a more criminological focus on crime indicative of a sociological criminology that fully encompasses “the processes of making laws, of breaking laws, and of reacting toward the breaking of laws (Sutherland and Cressey 1978:3).”
Bibliography


MN Statute § 611.17. 2021. *Financial Inquiry; Statements; Co-Payment; Standards for District Public Defense Eligibility*.


Walker, K. and Herman, M. 2022. tidycensus: Load US Census Boundary and Attribute Data as 'tidyverse' and 'sf'-Ready Data Frames. R package version 1.2.2. [https://CRAN.R-project.org/package=tidycensus](https://CRAN.R-project.org/package=tidycensus)


