

Appearance Discrimination in Criminal Court

By

Hannah Jean Frank

Dissertation

Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Law and Economics

May 10, 2019

Nashville, Tennessee

Approved:

Jennifer B. Shinall, Ph.D., J.D.

Kenworthy Bilz, Ph.D., J.D.

Nancy J. King, J.D.

Paige M. Skiba, Ph.D.

Copyright © 2019 by Hannah Jean Frank
All Rights Reserved

To the clients who have inspired me and motivated me,

To my amazing family,

and

To my partner (in bridge and in life)

ACKNOWLEDGEMENTS

This work would not have been possible without the incredibly generous educational and financial support of the Vanderbilt Law and Economics program. I feel very fortunate to have been a part of such a well-rounded and impeccably staffed program.

I am particularly grateful to all of the professors who served as members of my Dissertation Committee and as examples of smart, strong, and successful women in the academy. Thank you for elevating the quality of my scholarship. I am particularly indebted to my advisor, Dr. Jennifer Shinall, for all of her advice and guidance throughout the process. I could not have asked for a better mentor.

Throughout my time at Vanderbilt, I have been blessed with extraordinary peers. Thank you for reading my drafts, watching my presentations, and supporting me every step of the way. I would also like to thank all of the Vanderbilt professors who provided feedback on my projects and taught me how to research, write, and think critically.

Finally, a huge thank you to my family for your constant love and support. My parents taught me the value of education and the importance giving back, my sisters remind me every day how vital it is to be passionate about what you do, and my partner, Richard, has, and continues to be, the source of my greatest happiness.

TABLE OF CONTENTS

	Page
DEDICATION	iii
ACKNOWLEDGEMENTS.....	iv
LIST OF TABLES	vii
LIST OF FIGURES	x
INTRODUCTION	1
Chapter	
1. Race-Based Jury Nullification and Other Effects of Race in Nonviolent Criminal Cases	4
I. Introduction.....	4
II. Theoretical Foundations.....	8
III. Existing Literature	12
IV. Study Design.....	20
V. White Defendants Are Convicted More Often Than Black Defendants.....	29
VI. White Defendants Are Sentenced More Harshly Than Black Defendants When Mock Jurors Are Pooled.....	56
VII. Ruling Out the Attractiveness Alternative.....	70
VIII. Discussion and Conclusion.....	73
2. The Effects of Defendant Dress and Shackling on Nonviolent Criminal Case Outcomes	78
I. Introduction.....	78
II. Legal Background.....	80
III. Theoretical Foundations.....	84
IV. Existing Literature	88
V. Study Design.....	93
VI. Conviction Results.....	99
VII. Casual Attire Decreases Sentence Length and Institutional Attire Increases Sentence Length.....	126
VIII. Discussion and Conclusion.....	146
3. No Sympathy for the Shackled: Bans on Indiscriminate Juvenile Shackling Lead to More Favorable Case Outcomes for the Accused	153
I. Introduction.....	153
II. Literature Review.....	156

III. Background on Juvenile Justice in the United States	161
IV. Methodology	164
V. The Effect of Banning Indiscriminate Juvenile Shackling	186
VI. Heterogeneous Effects Across Crimes and Juveniles.....	196
VII. Discussion and Conclusion.....	200
APPENDIX	210
I. Appendix A: First Survey Outline	210
II. Appendix B: Second Survey Outline.....	222
APPENDIX TABLES	227
REFERENCES	230

LIST OF TABLES

Page

CHAPTER ONE

Table

1. Experimental Vignette Literature Examining the Effects of Race on Criminal Case Outcomes	17
2. First Vignette Study Subject Demographics.....	22
3. Second Vignette Study Subject Demographics	27
4. Conviction Rates by Defendant Race	29
5. Differences in White and Black Defendants' Conviction Rates by Mock Juror Race, Ethnicity, and Residence in the South Census Region	31
6. Demographic Characteristics of Actual and Prior Jurors	33
7. Differences in White and Black Defendants' Conviction Rates by Judicial and Juror Demographics and Prior Jury Service.....	34
8. Trial Evidence Believability by Defendant Race.....	36
9. Regression Analyses of Defendant Race and Evidence Believability on Conviction Rates	37
10. Regression Analyses of Defendant Race, Evidence Believability, and Race-Based Weighing of Evidence Believability on Conviction Rates	38
11. Pre-Evidence Assessments of Guilt by Defendant Race	40
12. Regression Analyses of Defendant Race and Beliefs About the Crime Charged, the Defendant's Propensity to Commit Crime, and the Criminal Justice System on Pre-Evidence Assessments of Guilt.....	43
13. Burden of Proof Interpretations by Defendant Race.....	49
14. Regression Analyses of Defendant Race, Perceived Importance of a Wrongful Conviction, Perceived Importance of a Wrongful Acquittal, and Perceived Importance of the Burden of Proof on Numerical Interpretations of the Burden of Proof	51
15. Regression Analyses of Defendant Race, Evidence Believability, Pre-Evidence Guilt Assessments, and Burden of Proof Interpretations on Conviction Rates	54
16. Sentence Length by Defendant Race	58
17. Disparities in Sentences Given to White and Black Defendants by Mock Juror Race, Ethnicity, and Residence in the South Census Region	59
18. Disparities in Sentences Given to White and Black Defendants by Judicial and Juror Demographics and Prior Jury Service.....	61
19. Disparities in Sentences Given to White and Black Defendants by Verdict Rendered....	63
20. Recidivism Predictions by Defendant Race.....	65
21. Racial Disparities in the Percent of Mock Jurors Predicting That the Defendant Will Recidivate by Verdict Rendered	66

22.	Regression Analyses of Defendant Race and Recidivism Predictions on Sentence Length by Verdict Rendered	68
23.	Regression Analyses of Defendant Race, Recidivism Predictions, and Race-Based Weighing of Recidivism Predictions on Sentence Length by Verdict Rendered	68
24.	Perceptions of Defendant Attractiveness by Defendant Race	70
25.	Regression Analyses of Defendant Race and Perceived Defendant Attractiveness on Case Outcomes	72

CHAPTER TWO

Table

1.	Subject Demographics	95
2.	Conviction Rates by Defendant Appearance	100
3.	Conviction Rates of White Defendants by Defendant Appearance.....	102
4.	Conviction Rates of Black Defendants by Defendant Appearance	102
5.	Regression Analyses of Defendant Dress and Shackling on Conviction Decisions Made by Democrat, Republican, Male, and Religious Mock Jurors and Mock Jurors Who Have No Criminal History	104
6.	Conviction Rates by Defendant Appearance, Judicial and Juror Demographics, and Prior Jury Service.....	107
7.	Regression Analyses of Defendant Dress and Shackling on the Believability of the Insurance Agent’s and the Defendant’s Testimony in Motive Vandalism Cases.....	114
8.	Pre-Evidence Assessments of Guilt by Defendant Appearance	117
9.	Regression Analyses of Defendant Dress and Shackling on Pre-Evidence Guilt Assessments of Republican, Male, and Religious Mock Jurors and Mock Jurors with No Criminal History	119
10.	Relative Importance of the Defendant’s Perceived Propensity to Commit Crime, the Crime Charged, and the Mock Juror’s Beliefs About the Criminal Justice System to the Presumption of Innocence Provided to Defendants in Institutional and Formal Attire by Republican, Male, and Religious Mock Jurors and Mock Jurors with No Criminal History	120
11.	Burden of Proof Interpretations by Defendant Appearance	122
12.	Relative Importance of Wrongful Conviction Costs, Wrongful Acquittal Costs, and the Burden of Proof to the Burden of Proof Applied to Defendants in Casual and Formal Attire	124
13.	Sentence Length by Defendant Appearance	130
14.	Sentencing of White Defendants by Defendant Appearance.....	131
15.	Sentencing of Black Defendants by Defendant Appearance	131
16.	Sentence Length by Defendant Appearance, Judicial and Juror Demographics, and Prior Jury Service.....	134
17.	Recidivism Predictions by Defendant Appearance.....	137
18.	Recidivism Predictions for White Defendants by Defendant Appearance.....	138
19.	Recidivism Predictions for Black Defendants by Defendant Appearance	138

20.	Regression Analyses of White Defendant Dress, Shackling, and Recidivism Predictions on Sentencing by Mock Jurors Who Convicted	140
21.	Regression Analyses of Defendant Dress, Shackling, and Perceived Attractiveness on Sentencing.....	144

CHAPTER THREE

Table

1.	State-Wide Bans on Indiscriminate Juvenile Shackling	154
2.	Juvenile Case Summary Statistics	180
3.	Crime Category Distributions in North Carolina, Tennessee, and Nationwide (in 2008)	183
4.	Double-Difference Estimates of the Effect of North Carolina’s Indiscriminate Juvenile Shackling Ban on Case Outcomes	188
5.	Double-Difference Estimates for Juveniles Ages 6-15.....	190
6.	Triple-Difference Estimates of the Effect of North Carolina’s Indiscriminate Juvenile Shackling Ban on Case Outcomes	192
7.	Triple-Difference Estimates for Juveniles Ages 6-15.....	195
8.	Triple-Difference Estimates for Various Subsets of Cases.....	197

APPENDIX TABLES

Table

A1.	First Study Subject Demographics by Race/Ethnicity	227
A2.	Second Study Subject Demographics by Race/Ethnicity	227
A3.	Case Demographics for Cases Involving Juveniles Ages 6–15	228

LIST OF FIGURES

Page

CHAPTER ONE

Figure

1.	Mechanisms Through Which Race May Impact Verdict Decisions.....	11
2.	Conviction Rates by Defendant Race	30
3.	Conviction Rates by Defendant Race and Mock Juror Race	32
4.	Presumption of Innocence by Defendant Race	41
5.	Burden of Proof by Defendant Race	50
6.	Distribution of Sentences Assigned in Vandalism Cases	57
7.	Distribution of Sentences Assigned in Arson Cases.....	58
8.	Sentence Length by Defendant Race and Verdict Rendered	64
9.	Recidivism Predictions by Defendant Race.....	66
10.	Perceived Attractiveness by Defendant Race	71

CHAPTER TWO

Figure

1.	Mechanisms Through Which Appearance May Impact Verdict Decisions.....	87
2.	Example Defendant Photographs.....	97
3.	Conviction Rates by Defendant Appearance	101
4.	Believability of Eyewitness Vandalism Trial Aspects by Defendant Appearance.....	110
5.	Believability of Eyewitness Arson Trial Aspects by Defendant Appearance	111
6.	Believability of Motive Vandalism Trial Aspects by Defendant Appearance	112
7.	Believability of Motive Arson Trial Aspects by Defendant Appearance	113
8.	Presumption of Innocence by Defendant Appearance.....	118
9.	Burden of Proof by Defendant Appearance.....	123
10.	Distribution of Sentences Assigned in Vandalism Cases	128
11.	Distribution of Sentences Assigned in Arson Cases.....	129
12.	Perceived Attractiveness by Defendant Appearance	143

CHAPTER THREE

Figure

1.	Juvenile Case Processing in the United States in 2016.....	164
2.	Trends in Case Outcomes	185
3.	Trends in Case Sentencing.....	186
4.	Double-Difference Results.....	189

5.	Age-Restricted Double-Difference Results	191
6.	Triple-Difference Results.....	194
7.	Age-Restricted Triple-Difference Results	196

INTRODUCTION

Perhaps nowhere is accuracy and precision more important than in a criminal courtroom. In the United States, 2,366 individuals have been exonerated after being convicted of crimes they did not commit and serving more than 20,735 years in prison (National Registry of Exonerations). Over the years, recognition of the enormity of the potential consequences of a criminal conviction has been the catalyst for a number of legal provisions designed to help ensure just outcomes in criminal cases. This dissertation examines the effects of various aspects of defendant appearance on criminal case outcomes to help further this important mission. The subtlety of appearance bias makes it particularly important to examine empirically and prevent ex-ante, as case-by-case litigations and ex-post corrections are likely to be plagued by evidentiary issues.

Chapter One begins by examining the effect of one of the most widely litigated and researched aspects of defendant appearance—race. Using an experimental vignette study involving two different nonviolent property crimes, I show that white defendants are convicted more often and receive longer prison sentences in these types of criminal trials. In part, this surprising result stems from race-based jury nullification. Mock jurors who distrust the justice system disproportionately acquit black defendants, and when mock jurors who would have acquitted a black defendant are asked to assign a sentence, they assign a disproportionately lesser sentence than similarly situated mock jurors who would have acquitted a white defendant. Additionally, despite the overarching results, there are racial stereotypes disadvantaging black defendants that affect both verdict and sentencing decisions. Mock jurors believe that black defendants have a higher propensity to commit crime, and black defendants are

disproportionately punished (in terms of longer prison sentences) for being perceived as likely to reoffend, receiving significantly longer sentences from mock jurors who convicted.

Chapter Two uses the same experimental vignette study to examine the impact of defendant dress and shackling. While these aspects of appearance have been virtually unexplored in the existing literature, they have both received attention from the Supreme Court. In *Deck v. Missouri*, the Court held that adult defendants cannot be shackled in front of a jury without a compelling state justification,¹ and in *Estelle v. Williams*, the Court held that defendants can never be forced to appear before a jury in identifiable prison garb.² At least in the context of trials for nonviolent property crimes, the results provide some support for the *Estelle* decision. Certain subgroups of mock jurors convict defendants in institutional garb at higher rates than defendants in civilian clothing. Additionally, defendants in institutional garb receive significantly longer prison sentences than defendants in casual civilian clothing. In contrast, the results indicate that shackling usually only affects case outcomes when considered in combination with defendant dress—defendants in casual civilian clothing are convicted the least and receive the shortest prison sentences, especially when they are not shackled.

Finally, Chapter Three provides the first empirical evidence on the effect of indiscriminate juvenile shackling on case outcomes. Since 2007, thirty-two states and the District of Columbia have put policies in place to limit the circumstances under which juveniles can be shackled in court, and many more states are currently considering bans on indiscriminate juvenile shackling. Using double-difference and triple-difference analyses and juvenile case data from North Carolina (where a ban on indiscriminate juvenile shackling was passed in 2007) and Tennessee (where no ban was passed until 2016) spanning 2004 through 2012, this Chapter

¹ 544 U.S. 622 (2005).

² 425 U.S. 501 (1976).

sheds light on the effects of such bans. The results indicate that North Carolina's ban on indiscriminate juvenile shackling resulted in more favorable outcomes for juvenile defendants, including a 3-4 percentage point decrease in the probability of receiving a sentence involving detention and an equally large increase in the probability of receiving a treatment-based sentence, providing support for a nationwide ban on indiscriminate juvenile shackling.

CHAPTER ONE: RACE-BASED JURY NULLIFICATION AND OTHER EFFECTS OF RACE IN NONVIOLENT CRIMINAL CASES

I. Introduction

The right of adult criminal defendants to have their cases resolved absent the influence of racial bias is protected by three federal constitutional provisions. First, the Due Process Clauses of the Fifth and Fourteenth Amendments guarantee the right to a fair trial, including the right to be presumed innocent and have one's guilt established by probative evidence beyond a reasonable doubt. Second, the Sixth Amendment states, "In all criminal prosecutions, the accused shall enjoy the right to ... an impartial jury."³ Third, the Equal Protection Clause of the Fourteenth Amendment promises defendants, "No state shall ... deny any person within its jurisdiction the equal protection of the laws." Similar protections are imposed on the federal government through the Due Process Clause of the Fifth Amendment.⁴

Despite these protections, racial discrimination at a criminal trial or sentencing proceeding is very hard to prove. In the case of a jury trial or jury sentencing proceeding, the defendant's first opportunity to discover racial bias by jurors is the jury selection process known as voir dire. During voir dire, if a potential juror admits to harboring racial prejudices or makes a statement illustrating a lack of impartiality, the defendant can move to have that juror dismissed for cause.⁵ However, courts often restrict the length and format of voir dire, limiting its effectiveness for identifying potential jurors with racial biases.⁶ Additionally, potential jurors are

³ The Supreme Court has also recognized the right to an impartial judge under the Due Process Clause. *Ward v. Village of Monroeville*, 409 U.S. 57 (1972).

⁴ *Bolling v. Sharpe*, 347 U.S. 497 (1954) (holding that the concept of equal protection is inherent in Due Process Clause of the Fifth Amendment).

⁵ The court should grant such a motion if the juror's views would "prevent or substantially impair the performance of his duties as a juror in accordance with his instructions and his oath." *Adams v. Texas*, 448 U.S. 38, 45 (1980).

⁶ For example, Federal Rule of Criminal Procedure 47 permits judges to choose whether the attorneys or the judge will ask questions during voir dire, thereby allowing judges to control the tone and delivery of inquiry.

likely to experience social pressure not to admit to being racially biased, and even the most honest potential jurors cannot admit to being implicitly biased, because, by definition, individuals are not consciously aware of their implicit biases.⁷

A defendant might also learn about racial bias after the jury has rendered a verdict. Since jury deliberations are private, this information usually comes in the form of an affidavit from one of the jurors. Unfortunately, such evidence may be inadmissible under no-impeachment rules, which limit the testimony that jurors can give about their deliberations in order to encourage vigorous deliberations and add finality to verdicts.⁸ Fearing that no-impeachment rules may be unduly restricting defendants' Sixth Amendment right to an impartial jury, the Supreme Court recently held that courts must consider juror affidavits reporting clearly racially biased statements made during jury deliberations.⁹ However, this is only the case when the statements tend to show that racial animus was a motivating factor in a juror's ultimate vote to convict.¹⁰

It may be even harder for a defendant to prove racial animus in the case of a bench trial or a judicial sentencing hearing. There is no voir dire process to determine which judge will hear the case, and since a judge makes his or her decision alone, there is no third party to report racially biased statements. Even if a defendant has evidence of racial animus, the judge being accused of bias will be the first to rule on the defendant's motion to have him or her replaced. If the motion is denied, the defendant will have to wait until he or she is convicted or sentenced to appeal, absent "exceptional circumstances."¹¹

⁷ For a discussion of the scientific foundations of implicit bias, see Greenwald and Krieger (2006); for papers discussing implicit bias in the criminal context, see Ingriselli (2015); Levinson, Cai, and Young (2010); Page (2005); and Rachlinski and Johnson (2009).

⁸ Every jurisdiction currently has a no-impeachment rule, but some are more restrictive than others. *See* Pena-Rodriguez v. Colorado, 137 S. Ct. 855, 865 (2017).

⁹ *Id.* at 869.

¹⁰ *Id.*

¹¹ *See* United States v. Gregory, 656 F.2d 1132 (5th Cir. 1981) (holding that the denial of defendant's motion to recuse was not immediately appealable under an exception to the typical rule that the defendant may only appeal

The fact that it is hard for defendants to prove racial prejudice on the part of a judge or jury could mean that constitutional violations are occurring without remedy. Empirically examining the effect of defendant race on case outcomes (and amending trial and sentencing procedures accordingly) can serve an important role in ensuring that all defendants are receiving their constitutionally guaranteed protections.¹² The fact that the vast majority of cases are now resolved through plea bargaining¹³ in no way diminishes the importance of studying and remedying racial prejudice at trial and sentencing. The bargaining process itself is influenced by the parties' expectations about what would happen at trial, and in cases where bargaining is affected by racial animus, trial may be the defendant's only opportunity for a fair adjudication.

A great deal of evidence already supports the concern that racial disparities exist within the criminal justice system. Black individuals, particularly young black men, are disproportionately represented at every stage of the criminal justice process (Blumstein 1993; Pope & Feyerherm 1995; Nellis 2016). In 2007, black individuals accounted for 38 percent of the United States jail and prison population and only 13 percent of the total population (Sentencing Project 2008). However, these observational disparities are not necessarily attributable to bias within the criminal justice system. They could reflect the fact that black individuals commit more crimes overall, that they commit more serious crimes, or that the

after a final judgment is entered in the case and that there were no exceptional circumstances justifying review under a writ of mandamus).

¹² General empirical evidence like that presented here will not allow an individual defendant to prove that their case outcome has been influenced by race *ex post*. *See* *McKleskey v. Kemp*, 481 U.S. 279 (1987) (holding that state-wide statistics on the race of death penalty defendants and their alleged victims could not prove intentional racial discrimination on the part of the local prosecutor). However, such evidence can motivate changes in policies that can help prevent the influence of race *ex ante*.

¹³ In 2014, 2.6 percent of federally adjudicated defendants received a bench or jury trial (Bureau of Justice Statistics 2017). In 2009, 2 percent of state felony convictions in the 75 largest counties in the United States resulted from a bench or jury trial (Bureau of Justice Statistics 2013).

system places heavy weight on race-neutral factors, like socioeconomic status or past encounters with the criminal justice system, which tend to correlate highly with race.¹⁴

In order to isolate the impact of defendant race on case outcomes, researchers often use experimental vignette studies that ask mock jurors to reach a verdict or recommend a sentence in a hypothetical criminal case. The results are mixed. Some studies have found no effect of defendant race (Bottoms, Davis, & Epstein 2004; Conley, Turnier, & Rose 2000; Hill & Pfeifer 1992; Hymes et al. 1993; McGuire & Bermant 1977; Nemeth & Sosis 1973; Sunnafrank & Fontes 1983), some studies have found bias against white defendants (Abwender & Hough 2001; Conley, Turnier, & Rose 2000; Skolnick & Shaw 1997; Sommers & Ellsworth 2000), and some studies have found bias against black defendants (Bernard 1979; Cohn et al. 2009; DeSantis & Kayson 1997; Foley & Chamblin 1982; Gordon 1993; Gray & Ashmore 1976; Ingriselli 2015; Pfeifer & Ogloff 1991; Sommers & Ellsworth 2000).

This Chapter will contribute to the existing literature using a pair of carefully designed experimental vignette studies to isolate the effect of defendant race on case outcomes in the context of a nonviolent property crime; the Chapter will further examine the source of any identified effect of defendant race. The results indicate that in the context of a trial and sentencing proceeding for a nonviolent property crime, decisionmakers convict white defendants significantly more often than black defendants and sentence white defendants significantly more harshly than black defendants. The bias against white defendants is at least partially the result of race-based jury nullification. Mock jurors' beliefs about biases present at other stages of the criminal justice process lead them to provide a stronger presumption of innocence to black

¹⁴ These explanations could in turn be explained by racial bias outside of the criminal justice system. For example, racial bias in employment could lead to socioeconomic disparities by race and make criminal behavior a more attractive alternative for black individuals.

defendants and perceive the evidence presented against black defendants as less persuasive. Additionally, mock jurors who acquitted the defendant sentence black defendants to significantly less time in prison than white defendants, possibly in an attempt to correct for what they perceive as racially-motivated convictions. Nevertheless, racial stereotypes disadvantaging black defendants influence both verdict and sentencing decisions. Mock jurors view black defendants as having a higher propensity to commit crime, and mock jurors who convicted the defendant sentence black defendants to longer prison terms despite viewing white defendants as more likely to recidivate.

II. Theoretical Foundations

A number of theoretical models have been developed to analyze and predict the decisionmaking process of judges and jurors, most of which have the same basic structure. A judge or juror reaching a verdict at trial in a criminal case should begin by presuming that the defendant is innocent. In other words, before any evidence is presented, the judge or juror should believe that there is a very small probability (approaching zero) that the defendant is guilty.¹⁵ As evidence is presented, the judge or juror weighs the credibility and sufficiency of the evidence and updates his or her belief about the defendant's likelihood of guilt. A number of mechanisms for this updating process have been proposed, perhaps the most widely accepted of which is often referred to as the story model (Pennington & Hastie 1991).¹⁶ Under this model, jurors assign meaning to trial evidence through the development of competing stories (Pennington &

¹⁵ Even absent racial bias, prior beliefs may stray from zero based on the difficulty of presuming innocence. For example, judges or jurors may improperly consider an indictment as evidence of guilt. A similar model of jury decisionmaking is discussed and tested in Ostrom et al. (1978).

¹⁶ Alternative models include information integration models, Bayesian models, Poisson models, sequential weighting models, and others (Pennington & Hastie 1981).

Hastie 1991). These stories are then judged based on pre-existing world knowledge to eliminate all but the “best” story (Pennington & Hastie 1991).

Whatever the precise mechanism for processing the information presented at trial, the judge or juror ultimately applies the jury instructions to their final belief about the likelihood of guilt in order to reach a verdict decision. There are two main ways to conceptualize this conviction decision task. First, a juror may translate the beyond a reasonable doubt standard of proof into a concrete percentage of guilt, such as 95 percent. If his or her post-evidence assessment of guilt is above that threshold, he or she will convict. Otherwise, he or she will not. Under this conceptualization, the threshold will be consistent across cases absent the influence of bias. Alternatively, the judge or juror might reach his or her verdict by weighing the costs of a wrongful conviction and wrongful acquittal. The weights that he or she assigns to each cost will depend on his or her post-evidence assessment of the defendant’s guilt and the size of the costs will depend on his or her personal beliefs and interpretation of the legal instructions. Under this conceptualization, the burden of proof threshold may vary between cases even absent bias. This is the conceptualization that best aligns with story model of jury decisionmaking. Under that model, the decisionmaker must decide whether the “best” story contains all the attributes necessary for a guilty verdict beyond a reasonable doubt (Pennington & Hastie 1991).¹⁷ This decision will depend on the decisionmakers prior ideas about crime categories, intuition, and interpretation of the legal instructions (Cheng 2013; Pennington & Hastie 1991).

¹⁷ Alternatively, jurors might convict if they believe that the prosecutor’s story is the most convincing and acquit otherwise, failing to consider the burden of proof in a third step. If this is the case, the burden of proof results presented here may not correlate with actual verdict decisions.

An unbiased judge or juror's pre-evidence belief, belief updating, and verdict determination will not be influenced by the defendant's race.¹⁸ In contrast, a biased judge or juror may—consciously or unconsciously—be influenced by the defendant's race in one of three different ways. First, before any evidence is presented, a biased judge or juror may allow racial stereotypes to influence his or her belief about the probability of the defendant's guilt. The judge or juror might associate members of a particular race with the crime charged, associate members of a particular race with criminal behavior more generally, or assign greater meaning to the charging of defendants of a particular race based on perceived biases in the system. Whatever the source of the stereotype, it will result in different initial probabilities of guilt for different defendants, in contravention of the presumption of innocence instruction.

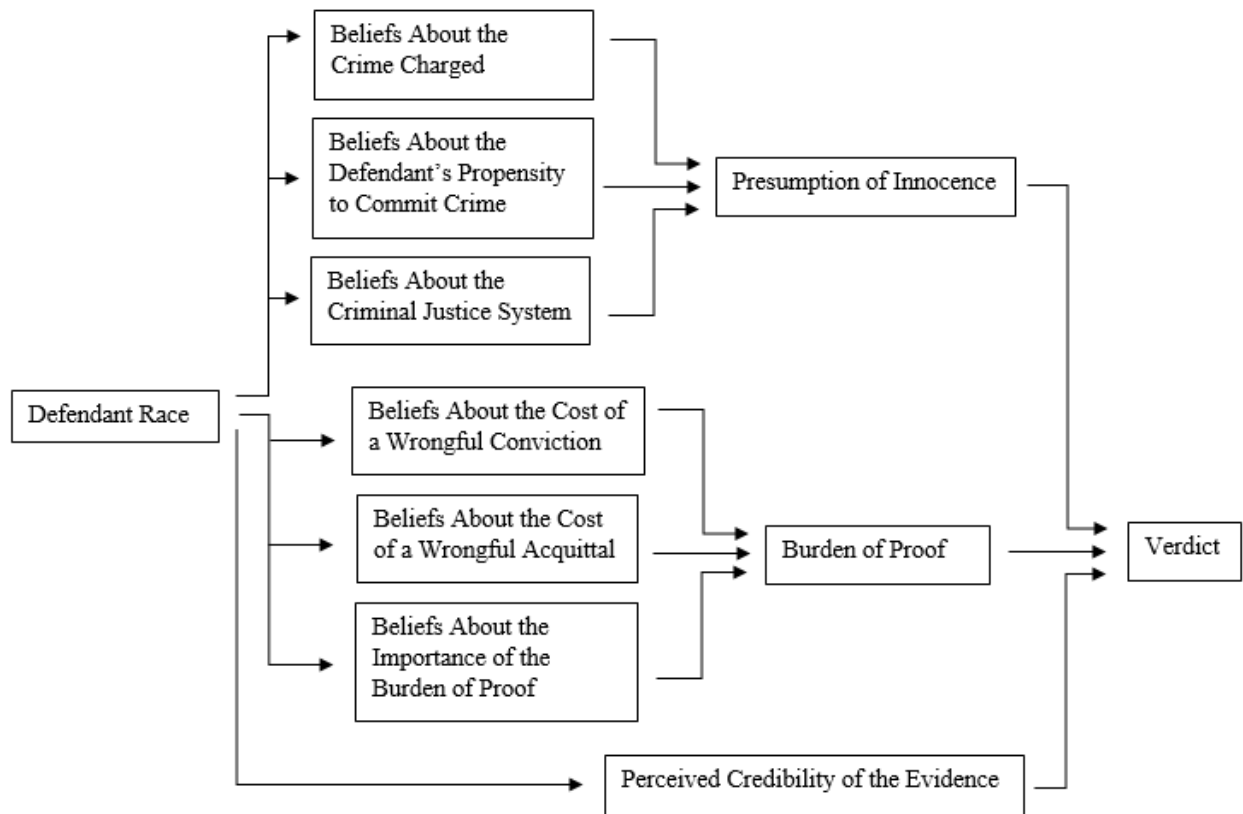
A biased judge or juror may also be influenced—consciously or unconsciously—by the defendant's race during the presentation of evidence. He or she may find certain types of evidence more credible based on the defendant's race or weigh the evidence differently. Finally, a biased judge or juror may—consciously or unconsciously—allow the defendant's race to affect his or her interpretation of the burden of proof. Beyond a reasonable doubt may be a more demanding standard for defendants of a certain race or the costs of a wrongful conviction and wrongful acquittal may differ based on the race of the defendant.

As Figure 1 illustrates, all of these forms of racial bias combine to influence aggregate differences in verdicts, and they may work in opposite directions. For example, it is possible that typical judges and jurors associate black individuals with criminal behavior and are therefore biased against black defendants even before any evidence is admitted. However, they may also find the evidence presented against white defendants more credible because they perceive no risk

¹⁸ Of course, even an unbiased judge or juror will consider race in cases where it is a relevant piece of evidence. Examples include hate crime cases and cases where the defendant claims bias on the part of police officers.

that the evidence has been influenced by racial prejudice. If they apply the same beyond a reasonable doubt standard in every case, it is not clear whether black or white defendants will be found guilty more often.

Figure 1. Mechanisms Through Which Race May Impact Verdict Decisions



Sentencing provides the judge or jury with much more discretion than the determination of guilt or innocence.¹⁹ The defendant has already been found guilty beyond a reasonable doubt, and the goals are to rehabilitate the defendant, deter future crime, incapacitate the defendant while he or she is a danger to the community, and obtain retribution. The judge or jury will

¹⁹ Most sentencing decisions are made by judges, but some sentencing decisions are left to juries, even outside the capital murder context. States that allow for jury sentencing in noncapital cases are Arkansas (ARK. CODE ANN. § 16-97-101), Kentucky (KY. REV. STAT. ANN. § 532.055), Missouri (MO. ANN. STAT. § 557.036), Oklahoma (OKLA. STAT. ANN. tit. 22 § 926.1), Texas (TEX. CODE CRIM. PROC. ANN. art. 37.07, § 2), and Virginia (VA. CODE ANN. § 19.2-295).

weigh these various goals based on their personal preferences and beliefs about the defendant and determine the optimal sentence within the statutorily permissible range. Sometimes, judges are provided with statutory sentencing guidelines to help them weigh various factors in coming to their decision.²⁰ Such guidelines are typically advisory and can be very complex.

An unbiased judge or juror will not consider the defendant's race in selecting the appropriate sentence. All else equal, black and white defendants will receive the same sentence for the same crime committed under similar circumstances. Conversely, a biased judge or juror may—consciously or unconsciously—think that defendants of a particular race are more capable of rehabilitation, more likely to reoffend, or more blameworthy. This will lead to aggregate differences in the sentences received by black and white defendants, absent any other differences in their cases or relevant personal circumstances.

III. Existing Literature

The generalized models of verdict determination and sentencing presented above lead to clear, testable hypotheses. If judges and jurors are influenced by defendant race, all else equal, defendants of a particular race will be convicted more often or sentenced more harshly than defendants of other races. Researchers have attempted to test these hypotheses using both real world observational data and experimental vignette studies. In both instances, the results are mixed.

²⁰ For a model of judicial sentencing incorporating sentencing guidelines and discussing the implications for racial sentencing disparities, see Sorenson et al. (2014).

A. Observational Studies

Many scholars have used data from real criminal cases to examine the effect of defendant race on case outcomes. Most of this research has used multivariate regression analysis to control for demographic characteristics and crime characteristics. Remaining racial disparities in case outcomes are typically presumed to be the result of racial bias. Primarily, these studies look at the effect of race on sentencing decisions (see Spohn 2000 for a review) and pretrial release and bail determinations (see Free 2001 for a review). Some studies have also looked at the effect of defendant race on dismissal decisions (Adams & Cutshall 1987; Myers 1982), decisions to seek the death penalty (Baldus, Woodworth, & Pulaski 1990; Nakell & Hardy 1987), and bindover decisions at preliminary hearings (Baumer, Messner, & Felson 2000).

Many of these studies find that after controlling for demographic and case characteristics, there is no effect of defendant race, indicating a lack of racial bias. Of the studies that do find evidence of racial bias, the majority find that black defendants are disadvantaged as compared to white defendants. However, these studies are all subject to the criticism that they may suffer from omitted variables bias because they can only control for the incomplete demographic and case characteristics that are observable in their data.

In order to avoid this criticism, a few researchers have used natural experiments or other unique methodological designs to analyze real world criminal case data. Alesina and Ferrara (2014) model trial courts as minimizing Type I (false positives, i.e. wrongfully sentencing a defendant to death) and Type II (false negatives, i.e. wrongfully failing to sentence a defendant to death) errors in the context of capital sentencing. This model leads to a testable hypothesis about racial bias—the rate at which death sentences are overturned during direct appeal and habeas corpus proceedings should not differ by the race of the defendant absent racial bias. The

authors find that the rate of reversal for death sentencing is higher for black defendants convicted of murdering white victims but only when those trials occur in the South. They conclude that trial judges in the South are racially biased against black defendants charged with murdering white victims. However, that conclusion relies on a few potentially problematic assumptions. First, the model assumes that appellate courts are not biased in favor of minority defendants when reviewing capital cases, which could be the case if appellate judges believe the outcomes of cases with minority defendants are more likely to be affected by racial bias. Second, the model assumes that there are no relevant unobservable case characteristics correlated with both defendant and victim race in the South that could serve as an alternative explanation for the different rates of reversal.

Abrams, Bertrand, and Mullainathan (2012) use the random assignment of cases to judges in Cook County, Illinois to examine whether defendant race impacts sentencing decisions. They find that the difference between a black defendant's and a white defendant's probability of incarceration when they are assigned to a judge who falls at the tenth percentile of judges in terms of racial disparities in incarceration decisions is eighteen percentage points greater than if they are assigned to a judge who falls at the ninetieth percentile of judges. They also find that the difference in the average incarceration sentence for those defendants increases by 10.47 percentage points when they move between the same two judges. A major limitation of their study is that they do not have a large enough sample to break down the results by crime type.

Natural experiments have also allowed researchers to examine a related question—how does the race of the judge or juror affect racial disparities in the criminal justice system? Anwar, Bayer, and Hjalmarsson (2012) use the random variation in jury pools to examine how the racial diversity of the jury pool interacts with defendant race to affect determinations of guilt. They

find that all-white jury pools convict black defendants sixteen percentage points more often than white defendants, and this gap is entirely eliminated when the jury pool contains at least one black individual. When they divide their sample by crime type, the effect goes away for property crimes. One limitation is that the study is somewhat confined in its scope. The data come exclusively from felony trials in two counties in Florida, each with a population of less than 500,000. Out of jury pools for 785 felony trials with an average size of 27 people, only 64 percent contained at least one black individual. As a result, the findings may only be representative of similarly small and non-diverse areas.²¹

Lee (2017) uses a difference-in-differences analysis to analyze the effect of policies that limit the discretion of jury commissioners to exclude black individuals from jury service. She finds that these reforms lead to a 5–6 percentage point drop in the share of new prison admissions accounted for by minority defendants. These results, while national in scope, could be driven by a particular type of case; the study focuses only on aggregate prison admissions, and the results are not broken down by type of crime charged. These natural experiments and related methodologies provide the most convincing results using real world data, but, as the discussion above suggests, each suffers from methodological and data-related limitations.

B. Experimental Studies

Many scholars have instead used experimental vignette studies to examine the effect of defendant race on criminal case outcomes. Typically, these studies ask jury-eligible subjects to read a written description of a criminal case or watch a video clip portraying a criminal case in which the race of the defendant is randomly assigned. Afterward, subjects are asked to determine

²¹ Florida is also unique in that it is one of only two states to allow for six-person juries in serious felony cases. FLA. R. CRIM. P. 3.720. The other is Connecticut. CONN. GEN. STAT. § 54-82.

guilt or recommend a sentence. This methodology has several benefits over observational studies. First, it allows the researcher to hold all case characteristics (other than defendant race) constant rather than having to measure and control for them to avoid omitted variables bias. Second, it allows researchers to easily collect data from a more demographically diverse sample. Finally, and perhaps most importantly, it allows researchers to examine the heterogeneous effect of defendant race across different types of cases, in terms of the crime charged, the evidence presented, and the courtroom procedures used.

Like studies using real-world case data, the results of these experimental studies are mixed. A number of experimental vignette studies find no effect of defendant race (Bottoms, Davis, & Epstein 2004; Conley, Turnier, & Rose 2000; Elek & Hannaford-Agor 2014; Hill & Pfeifer 1992; Hymes et al. 1993; McGuire & Bermant 1977; Nemeth & Sosis 1973; Sunnafrank & Fontes 1983), some find bias against white defendants (Abwender & Hough 2001; Conley, Turnier, & Rose 2000; Skolnick & Shaw 1997; Sommers & Ellsworth 2000), and some find bias against black defendants (Bernard 1979; Cohn et al. 2009; DeSantis & Kayson 1997; Foley & Chamblin 1982; Gordon 1993; Gray & Ashmore 1976; Ingriselli 2015; Pfeifer & Ogloff 1991; Sommers & Ellsworth 2000).

When studies find a bias against white defendants, it is typically concentrated among black respondents (Abwender & Hough 2001; Skolnick & Shaw 1997; Sommers & Ellsworth 2000). Studies have also found that the salience of race—measured by explicit discussion of race at trial—can increase racial bias (Cohn et al. 2009; Sommers & Ellsworth 2000), the provision of jury instructions can decrease racial bias (Hill & Pfeifer 1992; Mitchell et al. 2005), and the race of the victim can interact with the race of the defendant to influence racial bias (Bottoms, Davis, & Epstein 2004; Hymes et al. 1993). Table 1 summarizes the existing literature chronologically.

Table 1. Experimental Vignette Literature Examining the Effects of Race on Criminal Case Outcomes

Author(s)	Date of Publication	Crime	Mock Juror Sample Size	Outcome of Interest	Result
<i>Studies Finding Bias Against Black Defendants</i>					
Gray and Ashmore	1976	Manslaughter	380	Continuous sentencing determination	Bias against black defendants
Bernard	1979	Assault and battery of a police officer	115	Dichotomous assessment of guilt	Bias against black defendants
Field	1979	Rape	896	Continuous sentencing determination	Bias against black defendants
Foley and Chamblin	1982	Sexual battery on a child	191	Continuous assessment of guilt	Bias against black defendants by white participants
Pfeifer and Ogloff	1991	Rape	247	Dichotomous and continuous assessments of guilt	Bias against black defendants
Gordon	1993	Assault, embezzlement, and vehicular manslaughter	144	Continuous assessment of guilt	Bias against black defendants in assault cases
DeSantis and Kayson	1997	Burglary	160	Continuous sentencing determination	Bias against black defendants
Sommers and Ellsworth	2000	Assault, robbery, and arson	273	Continuous assessment of guilt and continuous sentencing determination	Bias against white defendants by black respondents and bias against black defendants by white respondents when race was salient
Cohn et al.	2009	Attempted vehicular homicide	176	Dichotomous assessment of guilt	Bias against black defendants when race was salient
Ingreselli	2015	Assault	412	Continuous assessment of guilt	Bias against black defendants
<i>Studies Finding No Racial Bias</i>					
Nemeth and Sosis	1973	Negligent vehicular homicide	80	Continuous sentencing determination	No significant effect of defendant race
McGuire and Bermant	1977	Murder	226	Dichotomous assessment of guilt	No significant effect of defendant race

Sunnafrank and Fontes	1983	Vehicular manslaughter	75	Dichotomous assessment of guilt	No significant effect of defendant race
Hill and Pfeifer	1992	Rape	125	Continuous assessment of guilt	No significant effect of defendant race once jury instructions were given
Hymes et al.	1993	Rape	78	Continuous assessment of guilt	No significant effect of defendant race
Conley, Turnier, and Rose	2000	Vandalism	551	Dichotomous assessment of guilt	Bias against white defendants that is eliminated in a follow-up study
Bottoms, Davis, and Epstein	2004	Sexual assault on a child	228	Dichotomous assessment of guilt	No significant effect of defendant race
Elek and Hannaford-Agor	2014	Assault and Battery	561	Dichotomous assessment of guilt and continuous sentencing determination	No significant effect of defendant race
<i><u>Studies Finding Bias Against White Defendants</u></i>					
Skolnick and Shaw	1997	Murder	213	Continuous assessment of guilt and continuous sentencing determination	Bias against white defendants by black participants
Conley, Turnier, and Rose	2000	Vandalism	551	Dichotomous assessment of guilt	Bias against white defendants that is eliminated in a follow-up study
Sommers and Ellsworth	2000	Assault, robbery, and arson	273	Continuous assessment of guilt and continuous sentencing determination	Bias against white defendants by black respondents and bias against black defendants by white respondents when race was salient
Abwender and Hough	2001	Negligent vehicular homicide	207	Continuous assessment of guilt and continuous sentencing determination	Bias against white defendants by black participants

These results illustrate the strength of the experimental vignette methodology—racial bias differs based on case-specific characteristics that are necessarily glossed over in studies

using real world criminal case data.²² However, existing experimental vignette studies are not without their own limitations. Most notably, researchers have thus far failed to meaningfully vary the crime type examined. Despite the fact that violent crimes account for only 5 percent of all arrests²³ and the fact that previous experimental vignette studies have illustrated that the race of the victim influences the effect of defendant race on case outcomes (Bottoms, Davis, & Epstein 2004; Hymes et al. 1993), only two experimental vignette studies have examined the effect of defendant race using a nonviolent crime. Importantly, there are reasons to suspect that the effect of race might be different in violent crime cases. Previous studies note that black defendants are often portrayed as more menacing and dangerous by the media (Dixon & Maddox 2005; Peffley, Shields, & Williams 1996), and a study by Hurwitz and Peffley (1997) finds that a strong relationship between whites' views on African-Americans and their assessments of a black defendant's likelihood of guilt (knowing only the race of the defendant and the crime charged) exists only when the crime charged is violent.

The two experimental studies that have not used a violent crime find either no racial bias or bias against white defendants. Gordon (1993) examines an embezzlement scenario. He finds no effect of defendant race on perceptions of guilt, but he does not provide any jury instructions, which have been found to influence racial bias (Hill & Pfeifer 1992; Mitchell et al. 2005). Conley, Turnier, and Rose (2000) examine a vandalism scenario and find mixed evidence, suggesting a possible bias against white defendants. Their experiment uses a group of released jury-pool members in an actual courtroom setting. The videotape of the trial includes jury

²² In some instances, these characteristics can be controlled for, but available samples of real world cases are typically not large enough for analyses of all potentially relevant subsamples of cases.

²³ This number was calculated using the Federal Bureau of Investigation's definition of violent crime, which includes murder and nonnegligent manslaughter, rape, robbery, and aggravated assault (Federal Bureau of Investigation 2017).

instructions, witness testimony, and attorney arguments, and the defendant is either black or white and either upper or lower class, as signaled by his clothing and his testimony about his education, career, and upbringing. Initially, the researchers found that the white defendant is convicted at a significantly higher rate than his black counterpart in both the upper- and lower-class versions of the study. In order to examine whether this effect is due to the fact that the black defendant was the only black individual in the tape, the authors then reran the study, substituting in black actors for some of the witnesses. In this second study, they find no significant effect of defendant race.

Additionally, Sommers and Ellsworth use a hypothetical arson scenario (along with a robbery scenario and an assault scenario) and find bias against white defendants by black mock jurors and bias against black defendants by white mock jurors when race is salient. However, their arson scenario could arguably be considered a violent hate crime, as it involves the burning down of a church with a congregation that is primarily of the opposite race of the defendant.

This study will contribute to the existing literature on the effect of defendant race on case outcomes using a pair of carefully designed experimental vignette studies involving nonviolent property crimes. This study will also examine how race is impacting case outcomes by examining the mechanisms illustrated in Figure 1.

IV. Study Design

This project was conducted using two experimental vignette studies. The main goals of the first study were to determine (1) whether defendant race significantly impacts conviction and sentencing results in cases involving nonviolent crimes, (2) whether the effects of defendant race differ based on the characteristics of the decisionmaker, (3) whether differences in mock jurors' perceptions of the evidence explain the effects of defendant race on verdict decisions (see Figure

1), and (4) whether differences in mock jurors' perceptions of the defendant's likelihood of recidivating explain the effects of defendant race on sentencing. A second study was conducted in order to (1) ensure that race, rather than differences in attractiveness, is the source of the racial disparities discovered in the first study, (2) determine whether differences in the presumption of innocence explain the effects of defendant race on verdict decisions (see Figure 1), and (3) determine whether differences in the burden of proof explain the effects of defendant race on verdict decisions (see Figure 1). The methodology used to conduct each of these experimental vignette studies is described in detail below.

A. First Study: Case-Specific Biases

The first experimental vignette study was conducted using Qualtrics survey technology and the Amazon Mechanical Turk (mTurk) survey platform. This platform provides access to a diverse nationwide sample of individuals who have signed up to take surveys for compensation. In total, 8,070 respondents participated in the survey, which included two separate vignette studies²⁴ and took about fifteen minutes to complete. Participants were only allowed to complete the survey once²⁵ and were compensated \$1.50 upon successful completion. All subjects were over eighteen, resided within the United States, and spoke English. Demographic characteristics of the sample are provided in Table 2 alongside U.S. population demographics from the 2016 Census. In general, the mTurk sample is somewhat younger, more educated, and more employed than the United States population. The mTurk sample also has a slight overrepresentation of women, an underrepresentation of racial and ethnic minorities, and is more urban than the United

²⁴ The other vignette study, which was not related to this project, involved employee hiring scenarios.

²⁵ This requirement is easily enforced since each participant must be signed up as an mTurk worker and each worker is assigned a unique worker ID.

States population.²⁶ Appendix Table A1 contains the first study subject demographics broken down by the race and ethnicity of subjects.

Table 2. First Vignette Study Subject Demographics

Characteristic	mTurk Sample	U.S. Population (2016)
Female	55.3%	51.1%
Age	35.4	41.8
US Native	95.0%	86.8%
Hispanic	8.1%	14.4%
White	79.0%	76.4%
Black	8.7%	10.2%
Asian	6.3%	5.4%
Married	41.1%	42.6%
B.A. or Higher (If 25+)	57.7%	31.4%
Employed	82.3%	46.2%
Median Household Income (\$2017)	\$52,500	\$58,844
Urban	33.7%	21.8%
Northeast	19.1%	17.7%
Midwest	21.3%	21.5%
South	38.5%	37.3%
West	20.9%	23.4%
Democrat	43.3%	N/A
Republican	22.4%	N/A
Religious	23.6%	N/A
Prior Jury Service	13.6%	N/A
Criminal History	17.0%	N/A
Criminal Victimization	41.4%	N/A
Distrusts Justice System	44.0%	N/A

Notes: Mean values for the mTurk sample are calculated using the entire sample of mock jurors from the first experimental vignette study. Mean values for the U.S. population are calculated using the 2016 U.S. Census Bureau ACS sample. Age is calculated using employed persons only; race is calculated using only individuals who report a single race. Individuals who attend church at least once a month are considered religious. Individuals who have a prior criminal conviction or have been arrested are considered to have a criminal history. Individuals are considered to distrust the justice system if they strongly disagreed or somewhat disagreed with the statement that the criminal justice system works well in the United States.

One concern with the use of mTurk workers is that they may be systematically different from the general population. While this is certainly a possibility, previous use of mTurk workers by a number of other scholars (Balganesh, Manta, & Wilkinson-Ryan 2014; Ingriselli 2015; Ginther et al. 2014; Hersch & Moran 2015; Hersch & Shinall 2017; Wilkinson-Ryan & Baron

²⁶ These demographic characteristics are typical of mTurk samples (see Ginther et al. 2014; Hersch & Shinall 2017).

2008), including some that have directly compared mTurk workers to other validated panels (Kuziemko et al. 2015), some who have examined juror decisionmaking in the criminal context (Ingriselli 2015; Ginther et al. 2014), and some who have examined bias against historically disadvantaged groups (Hersch & Moran 2015; Hersch & Shinall 2017) diminishes the concern. Ginther et al. (2014) uses an mTurk sample to examine assessments of culpability and the application of jury instructions in mock criminal cases. The study involves a follow-up experiment to an original study that uses a nationally representative sample. Using the mTurk sample does not change the overall pattern of responses by participants.

Each respondent in the first study was presented with one vandalism case scenario and one arson case scenario. As noted above, very few studies have examined the effect of defendant race on case outcomes in the context of a nonviolent crime. The vandalism and arson cases had two principal variations: an eyewitness variation and a motive variation. Two variations were included to ensure that the results were not driven by a particular aspect of the criminal scenario. Subjects were randomly assigned to see either the eyewitness or motive variation of the vandalism scenario, and then they saw the opposite variation of the arson scenario.

At the beginning of each scenario, subjects saw the following prompt:

You have been selected to serve on a jury in a criminal trial. The defendant is [**Michael Jones/ James Williams/ Joshua Thomas/ David Jackson**], pictured below as he appears in court on the day of trial. The sole charge against Mr. [**Jones/ Williams/ Thomas/ Jackson**] is [**vandalism/ arson**]. He has pleaded not guilty. After reading the trial material, you will be asked to decide whether the prosecution has proven that Mr. [**Jones/ Williams/ Thomas/ Jackson**] is guilty of [**vandalism/ arson**] beyond a reasonable doubt.

The surnames used are the four within the top 20 most common surnames from the 2000 Census that had the closest population-representative racial mix (U.S. Census Bureau a). First names were selected to be as familiar and race-neutral as possible.

In the eyewitness variation (whether vandalism or arson), subjects read a trial summary for a case in which an eyewitness description of the suspect's clothing and vehicle led to a neighborhood canvas by police; that canvas eventually turned up a matching suspect with relevant evidence in his home. The main defense was an alibi; two of the defendant's friends claimed to have been hanging out with him all night. The motive variation (whether vandalism or arson) revolved around a potential insurance fraud scheme. Evidence existed to suggest the defendant needed to pay off credit card debt, stood to gain money from the alleged crime, and relevant evidence linking the defendant to the crime was uncovered in the defendant's home. Here again, the defendant's main defense was an alibi; the defendant's wife claimed to have been home with him all night on the night of the crime. These scenarios were extensively pretested to ensure a mix of guilty and not guilty verdicts.

Most importantly for this experiment, each scenario was accompanied by a picture of either a black or white defendant. Pictures of two different white defendants and two different black defendants were used. The defendant's clothing and the presence or absence of shackles was also randomly varied. The effects of the clothing and shackles manipulations are discussed in Chapter Two. No respondent ever saw the same defendant in both scenarios.

Only two races were included in the study for better precision. There were already 24 different variations of the arson and vandalism questions (2 variations x 2 races x 2 shackling conditions x 3 clothing conditions), necessitating a large sample size. Including additional races or ethnicities would have increased the necessary sample size, the cost, and the complexity of the study significantly. Black individuals were selected for study because they experience the largest disparities in terms of their representation in the criminal justice system (Nellis 2016).

Differential treatment of members of other races and ethnicities, however, remains an important area for future research.

Each scenario concluded with jury instructions about the burden of proof, the standard of proof, the presumption of innocence, the elements of the crime charged, and the right to remain silent. These jury instructions were modeled primarily after Illinois's model criminal jury instructions for the crimes of arson and criminal damage to property. The elements of vandalism were: (1) that the defendant, (2) knowingly, (3) damaged property of another, (4) without the owner's consent. The elements of arson were: (1) that the defendant, (2) by means of fire or explosives, (3) knowingly, (4) damaged property of another, (5) without the owner's consent.

After viewing each scenario, respondents were asked to find the defendant either guilty or not guilty, rate their confidence in their guilt determination on a five point Likert scale, rate the believability of each witness and each party's version of events on a five point Likert scale, recommend a sentence, rate their confidence in their sentence determination on a five point Likert scale, and rate the likelihood that the defendant would reoffend in the next five years if acquitted on a five point Likert scale. The questions were asked in this order on separate screens, and respondents could not go back and change a previous answer.

For the sentencing determination, participants were asked to assume the role of a judge and, regardless of their answers to the previous questions, assume the defendant had been found guilty. They were told that the defendant's criminal history consisted of one misdemeanor theft conviction as a juvenile, and their answers were restricted to between one and five years (inclusive) for vandalism and between one and ten years (inclusive) for arson. These sentencing ranges are roughly representative of state sentencing ranges for these charges. To simplify the decision, there was not an option to impose a non-incarceration sentence, such as probation. At

the end of the survey, respondents were asked a variety of demographic questions. A complete survey outline is attached as Appendix A.

B. Second Study: Pre-Existing Biases

In order to explore whether pre-existing race-related beliefs influence the presumption of innocence and the burden of proof and to ensure that race rather than attractiveness is responsible for the racial disparities in the first study, a second study was conducted, again using Qualtrics survey technology and the mTurk survey platform. In total, 2,561 individuals participated, and they were compensated \$0.25 for about 2.5 minutes of their time. Again, all of the subjects were over eighteen, resided within the United States, and spoke English. Demographic characteristics of the sample are provided in Table 3, alongside those of the first mTurk sample and U.S. population demographics from the 2016 Census. For the most part, the demographic characteristics of the second sample mirror those of the first sample. One notable difference is a more representative proportion of Hispanic individuals (bringing the percent closer to that of the U.S. population). Additionally, as compared to the first study, a greater percentage of the second study respondents live in urban areas, identify as Republican, identify as religious, have prior jury experience, and trust the justice system. While these differences are large, they are not a major concern. The samples are drawn from the same population of mTurk workers, and—as discussed in more detail below—when the samples are analyzed together, the demographic characteristics of each sample and the results from the second study are used to predict values for the first sample, effectively controlling for any meaningful demographic differences between the samples. Appendix Table A2 contains the second study subject demographics broken down by the race and ethnicity of the subject.

Table 3. Second Vignette Study Subject Demographics

Characteristic	Second mTurk Sample	First mTurk Sample	U.S. Population (2016)
Female	54.5%	55.3%	51.1%
Age	35.3	35.4	41.8
US Native	94.4%	95.0%	86.8%
Hispanic	12.1%	8.1%	14.4%
White	79.2%	79.0%	76.4%
Black	8.5%	8.7%	10.2%
Asian	6.7%	6.3%	5.4%
Married	46.6%	41.1%	42.6%
B.A. or Higher (If 25+)	60.7%	57.7%	31.4%
Employed	84.9%	82.3%	46.2%
Median Household Income (\$2017)	\$52,500	\$52,500	\$58,844
Urban	39.5%	33.7%	21.8%
Northeast	17.0%	19.1%	17.7%
Midwest	19.3%	21.3%	21.5%
South	42.6%	38.5%	37.3%
West	20.9%	20.9%	23.4%
Democrat	40.2%	43.3%	N/A
Republican	30.5%	22.4%	N/A
Religious	34.7%	23.6%	N/A
Prior Jury Service	16.7%	13.6%	N/A
Criminal History	18.8%	17.0%	N/A
Criminal Victimization	41.9%	41.4%	N/A
Distrusts Justice System	31.9%	44.0%	N/A

Notes: Mean values for the first mTurk sample are calculated using the entire sample of mock jurors from the first experimental vignette study. Mean values for the second mTurk sample are calculated using the entire sample of mock jurors from the second experimental vignette study. Mean values for the U.S. population are calculated using the 2016 U.S. Census Bureau ACS sample. Age is calculated using employed persons only; race is calculated using only individuals who report a single race. Individuals who attend church at least once a month are considered religious. Individuals who have a prior criminal conviction or have been arrested are considered to have a criminal history. Individuals are considered to distrust the justice system if they strongly disagreed or somewhat disagreed with the statement that the criminal justice system works well in the United States.

Each respondent was presented with one criminal case scenario (either arson or vandalism) and then one attractiveness question. The criminal case scenarios began with the following prompt

You have been selected to serve on a jury in a criminal trial. The defendant is [**Michael Jones/ James Williams/ Joshua Thomas/ David Jackson**], pictured below as he appears in court on the day of trial. The sole charge against Mr. [**Jones/ Williams/ Thomas/ Jackson**] is [**vandalism/ arson**], and he has pleaded not guilty. At the end of the trial, you will be asked to apply the following rules of law:

1. Every person charged with a crime is presumed innocent.
2. You may decide what testimony to believe. You may believe all of the testimony of a witness, or part of it, or none of it.
3. The prosecution has the burden of proving every element of the crime beyond a reasonable doubt.
4. The elements of the crime of **[vandalism/ arson]** are:
 1. That the defendant,
 2. **[By means of fire or explosives,]**
 3. Knowingly,
 4. Damaged property of another,
 5. Without the owner's consent.

This prompt was accompanied by a randomly assigned picture of either a black or white defendant. The same pictures of the same two white and black defendants were used as in the first study. Participants were first asked how likely they thought it was that the defendant was guilty, on a scale of 0 to 100 percent, before seeing any evidence. Participants were then asked to rate the importance of (1) the crime charged, (2) the defendant's propensity to commit crime, and (3) their opinions about the criminal justice system to that rating of pre-evidence guilt on a five point Likert scale. Next, participants were asked how certain they would need to be of the defendant's guilt, on a scale of 0 to 100 percent, after hearing the evidence in order to return a guilty verdict. Participants were also asked to rate how important (1) fear of a wrongful conviction, (2) fear of a wrongful acquittal, and (3) whether the prosecution has met the burden of proof would be in deciding whether to convict on a five point Likert scale.

After the criminal case scenario, participants were asked to rate the attractiveness of a different defendant on a five point Likert scale. Like the criminal case scenarios, this question included a randomly assigned picture of a black or white defendant. The study concluded with a series of demographic questions. A complete survey outline is attached as Appendix B.

V. White Defendants Are Convicted More Often Than Black Defendants

In all of the case scenarios, white defendants are significantly more likely to be convicted than black defendants. Table 4 summarizes the conviction rates for each of the case scenarios in the first study and illustrates how they differ based on the race of the defendant. Arson cases result in higher conviction rates than vandalism cases and eyewitness variations result in higher conviction rates than motive variations. Eighty percent of participants are either somewhat confident or very confident that they have reached the legally correct verdict. Among those who convict the defendant, about 84 percent are somewhat or very confident about their decision, statistically significantly more than the approximately 77 percent of participants who feel the same after acquitting the defendant.

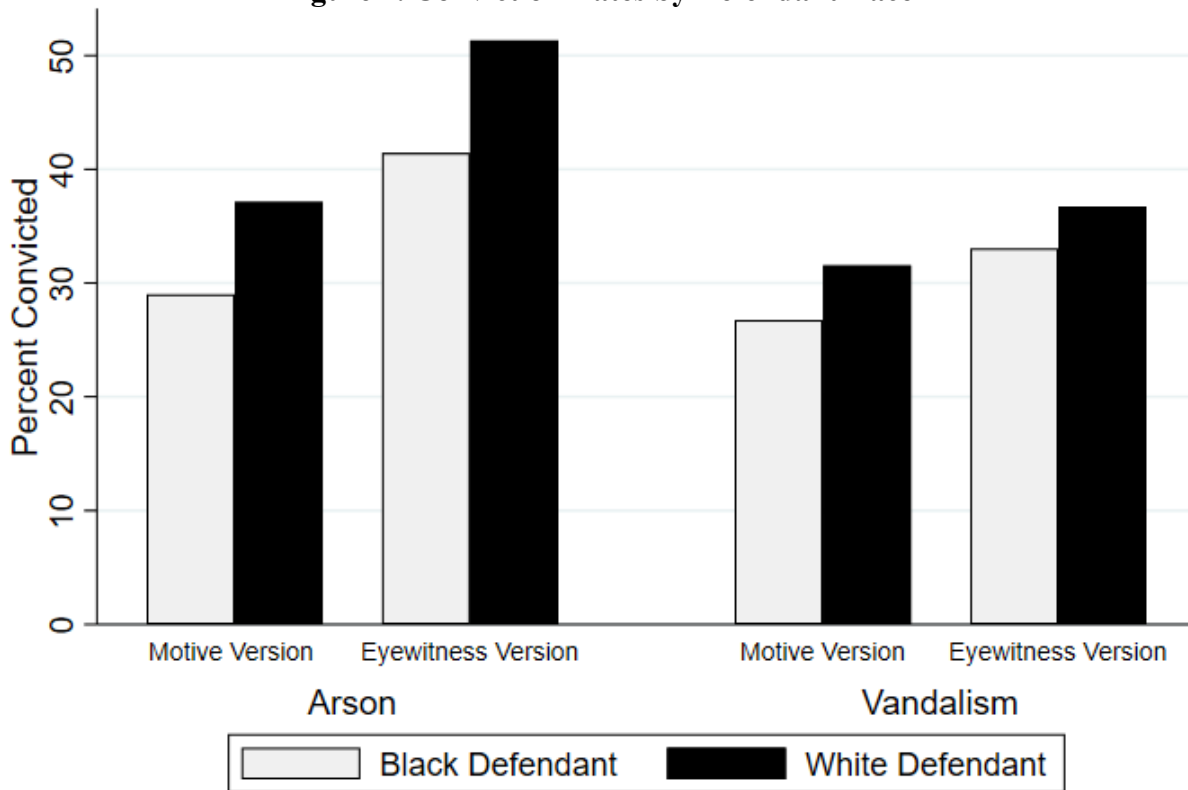
The biggest racial disparity occurs in eyewitness arson case scenario, where white defendants are about 10 percentage points more likely to be convicted than black defendants, and the smallest disparity occurs in eyewitness vandalism case scenario, where white defendants are about 4 percentage points more likely to be convicted than black defendants. Figure 2 illustrates average conviction probabilities by defendant race.

Table 4. Conviction Rates by Defendant Race

Case Scenario	All Defendants	White Defendants	Black Defendants
All Case Scenarios	35.82%	39.14%***	32.54%
Eyewitness Vandalism	34.94%	36.72%**	33.05%
Motive Vandalism	29.08%	31.54%***	26.76%
Eyewitness Arson	46.30%	51.33%***	41.46%
Motive Arson	33.06%	37.12%***	29.02%

Notes: Stars indicate significant differences between the conviction rates of black and white defendants. *p<0.1
p<0.05 *p<0.01.

Figure 2. Conviction Rates by Defendant Race



A. Mock Juror Heterogeneity: The Effect of Defendant Race Differs Based on the Race of the Mock Juror

Other experimental studies find that the effect of defendant race differs by the race of the mock juror (Abwender & Hough 2001; Skolnick & Shaw 1997; Sommers & Ellsworth 2000), and in an observational study, Alesina and Ferrara (2014) find that defendant race only has a significant effect on case outcomes in cases decided in the South. In order to test whether similar heterogeneity exists among mock jurors who participated in the first study, Table 5 contains the differences in white and black defendants' conviction rates by the race, ethnicity, and southern residency of the mock juror. While the racial disparities in each case scenario are not always significant, all of the subgroups of mock jurors convict white defendants significantly more often than black defendants when all of the case scenarios are pooled.

Table 5. Differences in White and Black Defendants' Conviction Rates by Mock Juror Race, Ethnicity, and Residence in the South Census Region

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson	Observations
White Only	4.77%***	2.45%	3.35%**	6.70%***	6.46%***	12,630
Black Only	16.92%***	9.44%*	17.98%***	28.36%***	14.09%***	1,396
Asian Only	11.25%***	9.88%	5.16%	12.16%**	18.60%***	1,012
Hispanic	7.82%***	5.40%	7.86%	9.11%	8.66%	1,268
Non-Hispanic	6.50%***	3.55%**	4.18%***	10.70%***	7.62%***	14,472
Southern	7.53%***	5.58%**	4.74%**	11.42%***	8.37%***	6,220
Not Southern	6.05%***	2.51%	4.88%***	8.96%***	7.93%***	9,920

Notes: Numbers represent conviction rates for white defendants minus conviction rates for black defendants rendered by respondents of the given race, ethnicity, or Census region in the given case scenario. Only individuals who report a single race are included in this analysis. Individuals are considered southern if their state of residence is in the south Census region and not southern otherwise. *p<0.1 **p<0.05 ***p<0.01.

In order to determine whether the magnitude of the bias against white defendants differs significantly by race, ethnicity, or southern residency, the following equation is estimated using a linear probability model for the subgroups of mock jurors listed in Table 5.

$$Guilty_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \epsilon_i \quad (1)$$

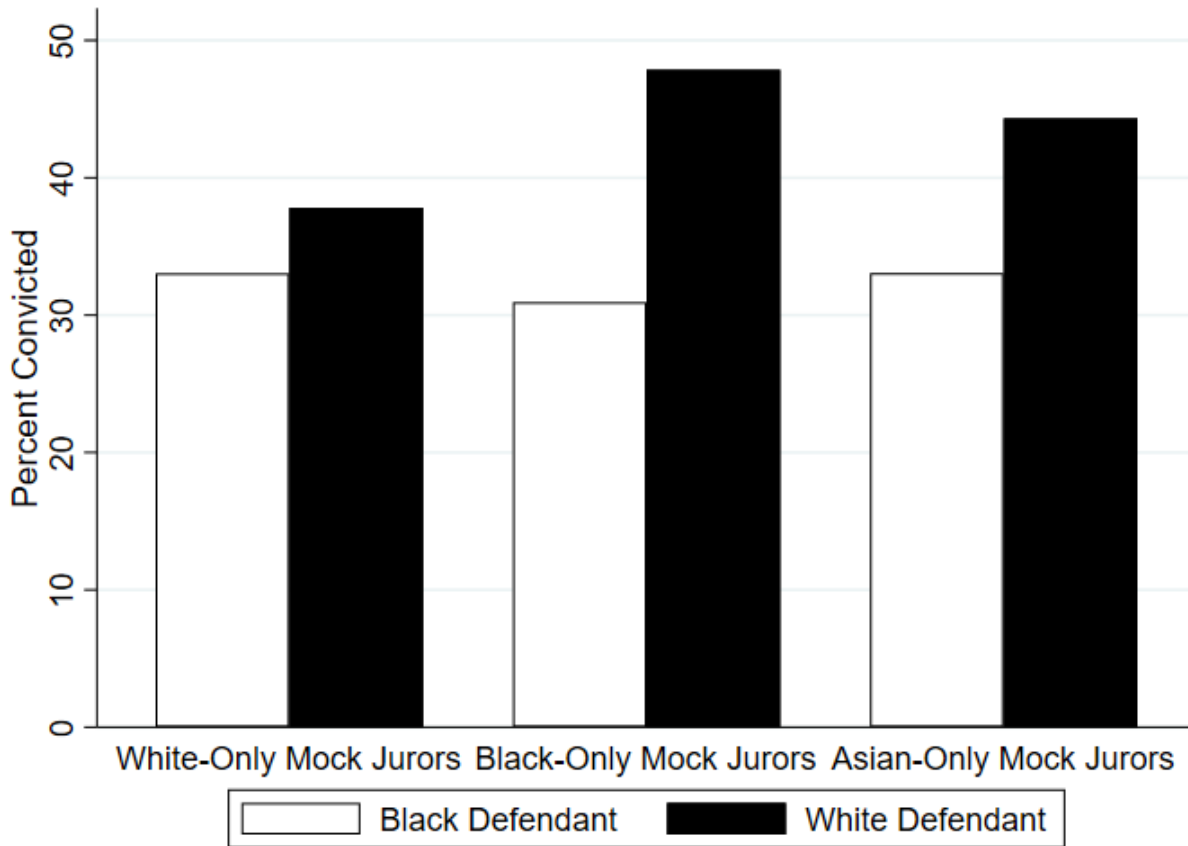
$Guilty_i$ is a dummy variable equal to one for a guilty verdict, $WhiteDefendant_i$ is a dummy variable equal to one when the defendant is white, $Arson_i$ is a dummy variable equal to one when the crime charged is arson, and $Eyewitness_i$ is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering around a motive. Standard errors are clustered at the individual level.

Once these regressions are estimated, the β_1 coefficients for related groups of mock jurors are tested for equivalency. The magnitude of the bias against white defendants differs significantly by the race of the mock juror, but it does not differ significantly by the ethnicity or southern residency of the mock juror.²⁷ As illustrated by the values in Table 5, black respondents

²⁷ Even though the magnitude of the race coefficient is not significantly different based on the ethnicity or southern residency of the mock juror, likelihood ratio tests indicate that including mock juror ethnicity and southern residency as explanatory variables in Equation 1 does significantly increase the fit of the model.

exhibit the strongest bias against white defendants, followed by Asian respondents and then white respondents. These results are depicted in Figure 3.

Figure 3. Conviction Rates by Defendant Race and Mock Juror Race



Even though the bias against white defendants appears robust—it is present in every case scenario and exists regardless of the race of the mock jurors—it is important to ensure that this bias is not driven by individuals in the sample who are unrepresentative of actual judges and jurors. According to the Bureau of Labor Statistics’ Occupational Employment Statistics survey, most, but not all, judicial positions require a law degree, and the median income for a judge or hearing officer in 2017 was \$115,520 per year (Bureau of Labor Statistics 2018).

Demographic characteristics of seated jurors are harder to come by, but Hannaford-Agor et al. (2003) collected demographic information on seated jurors via a juror survey in four courts

hearing non-capital felony cases during 2000 and 2001. The courts—the Los Angeles County Superior Court in California, Maricopa County Superior Court in Arizona, Bronx County Supreme Court in New York, and District of Columbia Superior Court in Washington, DC—were selected because each had a sufficiently high volume of felony trials to permit data collection within a reasonable period of time and court personnel were willing to cooperate with data collection procedures. The response rate among jurors was about 80 percent. The data were downloaded from the Inter-University Consortium for Political and Social Research.

Table 6 presents the demographic characteristics of seated jurors in the Hannaford-Agor et al. (2003) survey alongside U.S. population demographics from the 2016 Census and the demographic characteristics of individuals who report previous jury service in the first survey. Characteristics often differ between the Hannaford-Agor et al. (2003) sample and the mTurk prior juror sample,²⁸ but a few characteristics are consistent. In both samples of jurors, the highly educated are overrepresented, and the vast majority of individuals are employed.

Table 6. Demographic Characteristics of Actual and Prior Jurors

Characteristic	Hannaford-Agor et al. (2003) Actual Juror Sample	mTurk Prior Juror Sample	U.S. Population (2016)
Female	57.0%	54.1%	51.1%
Age	40.4	41.1	41.8
US Native	N/A	97.0%	86.8%
Hispanic	21.8%	7.7%	14.4%
White	59.0%	78.5%	76.4%
Black	26.4%	10.6%	10.2%
Asian	4.8%	5.3%	5.4%
Married	N/A	47.9%	42.6%
B.A. or Higher (If 25 or 26 +)	46.5%	62.1%	31.4%
Employed	85.7%	85.5%	46.2%

²⁸ There are particularly stark differences between the Hannaford-Agor et al. sample and the mTurk sample in terms of percent black, percent Hispanic, median household income, and religiosity. The differences in percent black and percent Hispanic might result from the fact that the Hannaford-Agor et al. sample is drawn only from urban populations, where minority populations are likely to be larger. The difference in religiosity might result from the different survey questions (see Table 6 notes). People who do not report attending church at least once a month may still identify as religious or very religious. The difference in income is harder to explain, especially given that the Hannaford-Agor et al. sample, which has a higher median income, is on average less educated.

Median Household Income (\$2017)	\$88,966	\$52,500	\$58,844
Urban	N/A	34.1%	21.8%
Northeast	23.0%	17.5%	17.7%
Midwest	0.0%	19.4%	21.5%
South	27.0%	40.0%	37.3%
West	50.0%	23.1%	23.4%
Democrat	N/A	41.1%	N/A
Republican	N/A	26.7%	N/A
Religious	65.7%	26.7%	N/A
Criminal History	N/A	19.6%	N/A
Criminal Victimization	N/A	53.3%	N/A
Distrusts Justice System	N/A	37.1%	N/A

Notes: Mean values for the U.S. population are calculated using the 2016 U.S. Census Bureau ACS sample. Age is calculated for employed persons only; race values are calculated for individuals who report only a single race. B.A. or higher is calculated for people 25+ in the mTurk prior juror sample and the U.S. population sample and for people 26+ in the Hannaford-Agor et al. (2003) sample. In the mTurk prior juror sample, individuals who attend church at least once a month are considered religious, and in the Hannaford-Agor et al. (2003) sample, individuals who reported being “religious” or “very religious” rather than “neither,” “somewhat nonreligious,” or “very nonreligious” are considered religious. Individuals who have a prior criminal conviction or have been arrested are considered to have a criminal history. Individuals are considered to distrust the justice system if they strongly disagree or somewhat disagree with the statement that the criminal justice system works well in the United States.

Table 7 uses this demographic information to present the differences in conviction rates between white and black defendants by those with and without the education level and approximate income level of judges, those with and without the most common characteristics of seated jurors, and those who have and have not previously served on a jury. Individuals are considered to have the education level and approximate income level of judges if they have a doctoral or professional degree and an income above \$75,000 per year. Individuals are considered to have the most common characteristics of seated jurors if they have at least some college education and are employed.

Table 7. Differences in White and Black Defendants’ Conviction Rates by Judicial and Juror Demographics and Prior Jury Service

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson	Observations
Prior Jury Service	6.83%***	3.04%	11.04%***	7.27%*	6.31%	2,200
No Prior Jury Service	6.57%***	3.77%**	3.71%**	10.31%***	8.38%***	13,940

Characteristics Similar to Seated Jurors	6.30%***	3.71%**	4.40%***	9.68%***	7.39%***	12,260
Characteristics Distinct from Seated Jurors	7.59%***	3.69%	5.99%**	10.55%***	10.20%***	3,880
Education and Income Similar to Judges	5.31%*	6.11%	7.11%	3.91%	3.85%	1,186
Education or Income Distinct from Judges	6.70%***	3.58%**	4.55%***	10.39%***	8.40%***	14,880

Notes: Numbers represent conviction rates for white defendants minus conviction rates for black defendants rendered by respondents with the listed characteristic in the given case scenario. Individuals are considered to have education and income similar to a judge if they have a doctoral or professional degree and if their income is \$75,000 per year or more. Individuals are considered to have characteristics similar to seated jurors if they have at least some college education and are employed. Mock jurors who elect not to report their income are not included in either the group of mock jurors with education and income similar to judges or the group of mock jurors with education and income distinct from judges. *p<0.1 **p<0.05 ***p<0.01.

Those with education and income levels similar to judges do not convict white and black defendants at significantly different rates in any individual case scenario. However, this lack of significance could be due to reduced sample sizes. This is especially likely given that the magnitude of the differences in conviction rates by race are still large and positive in each case scenario, and there is a marginally significant bias against white defendants when all of the case scenarios are combined. No clear pattern emerges in terms of prior jury service. While those with prior jury service exhibit significant racial biases less often than those without prior jury service, they exhibit a significant bias against white defendants in the motive vandalism case scenario that is larger than the bias against white defendants exhibited by those without prior jury service. Those with common juror characteristics exhibit slightly less of a bias against white defendants than those without these characteristics, but the bias is usually significant in both samples. These results constitute minimal evidence that the jury selection process and background experiences of judges help to decrease racial bias. Juror bias is never completely eliminated, and in some cases the bias against white defendants may actually be larger among seated jurors and judges.

B. Case-Specific Biases: Defendant Race Affects the Perceived Credibility of the Evidence and the Weight Assigned to Evidence Credibility

At least some of the racial bias exhibited by mock jurors may be case-specific in terms of being significantly correlated with the believability of the evidence and attorney arguments. Recall that after rendering a verdict, participants in the first study were asked to rate the believability of each witness and each party’s version of events. As we might expect from the conviction rates, significantly fewer participants find the prosecution’s version of events and the testimony of the prosecution’s witnesses very believable or somewhat believable when the defendant is black. Paralleling these results, significantly more participants find the defense’s version of events and the testimony of the defense’s witnesses very believable or somewhat believable when the defendant is black. Table 8 summarizes participants’ assessments of these aspects of the trial by the race of the defendant.

Table 8. Trial Evidence Believability by Defendant Race

Scenario	All Defendants	White Defendants	Black Defendants
<i><u>Eyewitness Vandalism</u></i>			
Testimony of Eyewitness	69.88%	71.12%*	68.57%
Testimony of Police Officer	75.96%	77.25%**	74.38%
Testimony of Defendant’s Friends	46.66%	43.34%***	50.18%
Prosecution’s Version of Events	75.52%	77.83%***	73.07%
Defense’s Version of Events	57.93%	55.33%***	60.69%
<i><u>Motive Vandalism</u></i>			
Testimony of Police Officer	75.83%	78.13%***	73.67%
Testimony of Insurance Agent	75.28%	76.53%*	74.11%
Testimony of Defendant’s Wife	51.64%	47.36%***	55.67%
Testimony of Defendant	51.29%	46.85%***	55.47%
Prosecution’s Version of Events	70.74%	74.25%***	67.45%
Defense’s Version of Events	56.93%	53.83%***	59.85%
<i><u>Eyewitness Arson</u></i>			
Testimony of Eyewitness	76.69%	79.78%***	73.71%
Testimony of Police Officer	80.77%	83.25%***	78.19%
Testimony of Defendant’s Friends	44.87%	40.96%***	48.65%
Prosecution’s Version of Events	78.54%	81.26%***	75.92%
Defense’s Version of Events	54.12%	50.92%***	57.21%
<i><u>Motive Arson</u></i>			
Testimony of Police Officer	75.28%	76.91%**	73.66%
Testimony of Insurance Agent	76.18%	78.29%***	74.10%

Testimony of Defendant's Wife	54.08%	51.50%***	56.63%
Testimony of Defendant	54.47%	51.35%***	57.56%
Prosecution's Version of Events	73.12%	75.28%***	70.98%
Defense's Version of Events	61.41%	58.54%***	64.24%

Notes: Respondents are considered to find an aspect of the trial believable if they said it was very believable or somewhat believable on a five point Likert scale. Stars indicate significant differences in the percentages of respondents who found the aspect of the trial believable for black and white defendants. *p<0.1 **p<0.05 ***p<0.01.

In order to determine whether the differences in perceived credibility of the witnesses and case theories account for any of the disparity in conviction rates between white and black defendants, the following equations are estimated using linear probability models.

$$Guilty_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 CredibilityRatings_i + \epsilon_i \quad (2)$$

$$Guilty_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 CredibilityRatings_i + \beta_3 WhiteDefendant_i * CredibilityRatings_i + \epsilon_i \quad (3)$$

$Guilty_i$ is a dummy variable equal to one for a guilty verdict, $WhiteDefendant_i$ is a dummy variable equal to one when the defendant is white, $CredibilityRatings_i$ is a vector including dummy variables equal to one if the respondent found each witness and theory of the case very believable or somewhat believable in the particular case scenario, and $WhiteDefendant_i * CredibilityRatings_i$ is a vector including interactions between $WhiteDefendant_i$ and $CredibilityRatings_i$. Table 9 contains the regression results for Equation 2, and Table 10 contains the regression results for Equation 3.

Table 9. Regression Analyses of Defendant Race and Evidence Believability on Conviction Rates

	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
White Defendant	-0.01	-0.001	0.04***	0.04***
Prosecutor's Version of Events	0.12***	0.13***	0.19***	0.17***
Defense's Version of Events	-0.33***	-0.23***	-0.32***	-0.26***
Testimony of Eyewitness	0.13***	—	0.13***	—
Testimony of Insurance Agent	—	0.03	—	0.05***
Testimony of Police Officer	0.09***	0.04**	0.05**	0.07***
Testimony of Defendant's Friends	-0.24***	—	-0.24***	—
Testimony of Defendant's Wife	—	-0.11***	—	-0.10***

Testimony of Defendant	—	-0.16****	—	-0.19***
Observations	4,081	3,989	3,989	4,081
R-squared	0.331	0.273	0.357	0.316

Notes: Standard ordinary least squares regressions are used. The outcome variable is a dummy variable equal to one if the defendant is found guilty. White defendant is a dummy variable equal to one when the defendant is white. Each of the other variables is a dummy variable equal to one if the participant rates that aspect of trial very believable or somewhat believable on a five point Likert scale. Standard errors are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

Table 10. Regression Analyses of Defendant Race, Evidence Believability, and Race-Based Weighing of Evidence Believability on Conviction Rates

	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
White Defendant	0.03	0.01	-0.04	0.02
Prosecutor's Version of Events	0.15***	0.12***	0.16***	0.16***
Defense's Version of Events	-0.34***	-0.20***	-0.31***	-0.27***
Testimony of Eyewitness	0.14***	—	0.13***	—
Testimony of Insurance Agent	—	0.03	—	0.08***
Testimony of Police Officer	0.06**	0.05*	0.04	0.04*
Testimony of Defendant's Friends	-0.21***	—	-0.30***	—
Testimony of Defendant's Wife	—	-0.09**	—	-0.06
Testimony of Defendant	—	-0.20***	—	-0.23***
White Defendant * Prosecutor's Version of Events	-0.07**	0.02	0.05	0.03
White Defendant * Defense's Version of Events	0.03	-0.05	-0.01	0.02
White Defendant * Testimony of Eyewitness	-0.02	—	0.002	—
White Defendant * Testimony of Insurance Agent	—	0.02	—	-0.02
White Defendant * Testimony of Police Officer	0.06	-0.04	0.02	0.01
White Defendant * Testimony of Defendant's Friends	-0.05	—	0.06	—
White Defendant * Testimony of Defendant's Wife	—	-0.04	—	-0.09
White Defendant * Testimony of Defendant	—	0.08	—	0.07
Observations	4,081	3,989	3,989	4,081
R-Squared	0.332	0.274	0.359	0.317

Notes: Standard ordinary least squares regressions are used. The outcome variable is a dummy variable equal to one if the defendant is found guilty. White defendant is a dummy variable equal to one when the defendant is white. Each of the other variables is either a dummy variable equal to one if the participant rates that aspect of trial very believable or somewhat believable on a five point Likert scale or an interaction of two prior variables. Standard errors are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

The results in Table 9 illustrate that controlling for perceived differences in the credibility of the evidence eliminates the significant effect of defendant race on conviction rates in

vandalism cases. Some significant effect of defendant race remains in the arson cases. However, when interaction effects are included to allow perceived credibility to affect case outcomes differently depending on the race of the defendant (see Table 10), defendant race fails to have a significant effect in any of the case scenarios. This indicates that the bias against white defendants in terms of conviction rates is primarily accounted for by race-based differences in perceptions of the evidence and race-based differences in the weight that those perceptions are given when deciding whether to convict. It is important to recognize that the evidence does not vary based on the race of the defendant, so the different perceptions of the credibility of the evidence that account for the differences in conviction rates are entirely race-based. Furthermore, it is still possible that pre-existing race-related beliefs also account for some of racial disparities in conviction rates. The next Part explores this possibility using the data from the second study.

C. Pre-Existing Biases

As illustrated in Figure 1, theory provides three major avenues through which the race of the defendant could influence verdict determinations—by affecting the presumption of innocence, by affecting perceptions of the evidence presented at trial, and by affecting the application of the burden of proof. The results in the previous Part illustrate that race influences perceptions of the evidence and weighing of evidence credibility such that controlling for mock jurors' perceptions of the evidence and the interaction between these perceptions and defendant race eliminates the significant effect of defendant race on convictions. This Part uses data from the second study to examine the influence of pre-existing biases on the presumption of innocence and the burden of proof.

1. White Defendants Are Presumed Less Innocent Than Black Defendants

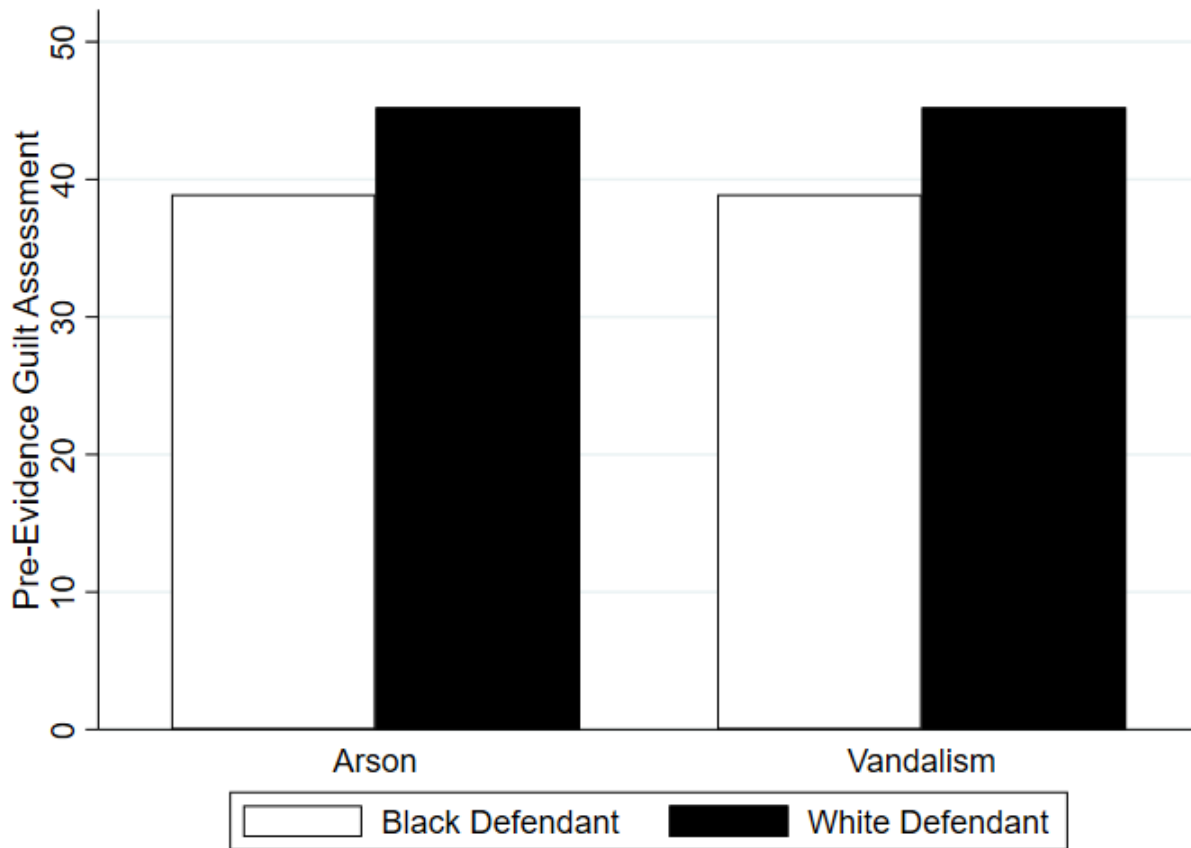
Recall that in the second study participants were asked how likely they thought it was that the defendant was guilty before seeing any evidence. Consistent with the conviction and evidence credibility results, regardless of the crime charged, respondents believe white defendants are significantly more likely to be guilty. The difference in the pre-evidence assessments of guilt assigned to black and white defendants is consistently between 6 and 7 percentage points. Table 11 presents and Figure 4 depicts the presumption of innocence results from the second study by defendant race.

Table 11. Pre-Evidence Assessments of Guilt by Defendant Race

Case Scenario	All Defendants	White Defendants	Black Defendants
All Cases	42.01%	45.25%***	38.92%
Vandalism	42.73%	45.91%***	39.66%
Arson	41.31%	44.59%***	38.18%

Notes: Numbers represent average estimates of the defendant's guilt on a scale from 0 to 100, inclusive, based only on a picture of the defendant, the crime charged, and the jury instructions. Stars indicate significant differences between the pre-evidence guilt assessments for black and white defendants. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

Figure 4. Presumption of Innocence by Defendant Race



As Figure 1 illustrates, these differences in pre-evidence assessments of guilt by defendant race could be the result of beliefs about the crime charged—for example, if respondents believe that vandalism and arson are crimes committed primarily by white defendants; beliefs about the association of race and criminal activity generally—for example, if respondents think that white defendants are more prone to committing crime; or beliefs about racial biases affecting other aspects of the criminal justice process—for example, if respondents believe that police officers or prosecutors exhibit racial bias against black individuals and they want to counteract those biases.

In order to explore these possibilities, the following equations are estimated using linear regression models. For these analyses, all of the cases are combined.

$$PresumptionofInnocence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 FactorImportance_i + \beta_3 Arson_i + \beta_4 DistrustsSystem_i + \epsilon_i \quad (4)$$

$$PresumptionofInnocence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 FactorImportance_i + \beta_3 WhiteDefendant_i * FactorImportance_i + \beta_4 Arson_i + \beta_5 Arson_i * CrimeImportant_i + \beta_6 Arson_i * WhiteDefendant_i + \beta_7 DistrustsSystem_i + \beta_8 DistrustsSystem_i * CrimJustBeliefsImportant_i + \beta_9 DistrustsSystem_i * WhiteDefendant_i + \epsilon_i \quad (5)$$

$$PresumptionofInnocence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 FactorImportance_i + \beta_3 WhiteDefendant_i * FactorImportance_i + \beta_4 Arson_i + \beta_5 Arson_i * CrimeImportant_i + \beta_6 Arson_i * WhiteDefendant_i + \beta_7 Arson_i * CrimeImportant_i * WhiteDefendant_i + \beta_8 DistrustsSystem_i + \beta_9 DistrustsSystem_i * CrimJustBeliefsImportant_i + \beta_{10} DistrustsSystem_i * WhiteDefendant_i + \beta_{11} DistrustsSystem_i * CrimJustBeliefsImportant_i * WhiteDefendant_i + \epsilon_i \quad (6)$$

PresumptionofInnocence_i is a variable equal to the respondents' pre-evidence guilt assessment. *FactorImportance_i* is a vector that includes three dummy variables equal to one if the respondent said the crime charged, the defendant's propensity to commit crime, and the respondent's beliefs about the criminal justice system are very or somewhat important on a five point Likert scale, respectively. *WhiteDefendant_i * FactorImportance_i* is a vector that contains the interactions of *WhiteDefendant_i* and *FactorImportance_i*. *Arson_i* is a dummy variable equal to one when the crime charged is arson. *Arson_i * CrimeImportant_i* is the interaction between *Arson_i* and a dummy variable equal to one if the respondent said that the crime charged is somewhat or very important. *Arson_i * WhiteDefendant_i* is the interaction between *Arson_i* and *WhiteDefendant_i*. *DistrustsSystem_i* is a variable equal to one if the

respondent strongly or somewhat disagrees with the statement that the criminal justice system works well in the United States. $DistrustsSystem_i * CrimJustBeliefsImportant_i$ is the interaction of $DistrustsSystem_i$ and a dummy variable equal to one if the respondent said that their beliefs about the criminal justice system are somewhat or very important.

$DistrustsSystem_i * WhiteDefendant_i$ is the interaction between $DistrustsSystem_i$ and $WhiteDefendant_i$. Finally, $DistrustsSystem_i * CrimJustBeliefsImportant_i * WhiteDefendant_i$ and $Arson_i * CrimeImportant_i * WhiteDefendant_i$ are triple-interactions between $DistrustsSystem_i$, $CrimeJustBeliefsImportant_i$, and $WhiteDefendant_i$, and $Arson_i$, $CrimeImportant_i$, and $WhiteDefendant_i$, respectively.

Of course, it is possible that respondents are not aware of, or do not want to admit, how important each of these factors is to their decision; these regressions only capture the conscious, reported effect of these various factors. Regression results are presented in Table 12.

Table 12. Regression Analyses of Defendant Race and Beliefs About the Crime Charged, the Defendant’s Propensity to Commit Crime, and the Criminal Justice System on Pre-Evidence Assessments of Guilt

	Equation 4	Equation 5	Equation 6
White Defendant	6.02***	2.99	5.83**
Crime Important	-0.74	-2.85	-1.64
Propensity Important	3.74***	5.77***	5.70***
System Beliefs Important	1.97*	2.22	4.05**
White Defendant * Crime Important	—	4.27**	1.76
White Defendant * Propensity Important	—	-4.53**	-4.52**
White Defendant * System Beliefs Important	—	-0.74	-4.55*
Distrusts System	-6.97***	-11.50***	-9.38***
Distrusts System * System Beliefs Important	—	0.52	-5.03*
Distrusts System * White Defendant	—	8.98***	4.53*
Distrusts System * System Beliefs Important * White Defendant	—	—	11.72***
Arson	-1.43	-1.65	-0.14
Arson * Crime Important	—	-0.02	-2.55
Arson * White Defendant	—	0.40	-2.82
Arson * Crime Important * White Defendant	—	—	5.45
Observations	2,512	2,512	2,512
R-Squared	0.042	0.052	0.055

Notes: Standard ordinary least squares regressions are used. The outcome variable is the respondent's estimate of the defendant's guilt on a scale from 0 to 100, inclusive, based only on a picture of the defendant, the crime charged, and the jury instructions. White defendant is a dummy variable equal to one when the defendant is white. Crime Important is a dummy variable equal to one if the respondent said the crime charged is very important or somewhat important in determining pre-evidence guilt on a five point Likert scale. Propensity important is a dummy variable equal to one if the respondent said the defendant's propensity to commit crime is very important or somewhat important in determining pre-evidence guilt on a five point Likert scale. System Beliefs Important is a dummy variable equal to one if the respondent said that their beliefs about the criminal justice system are very important or somewhat important in determining pre-evidence guilt on a five point Likert scale. Distrusts System is a dummy variable equal to one if the respondent said they strongly or somewhat disagree that the justice system works well in the United States on a five point Likert scale. Arson is a dummy variable equal to one if the crime charged is arson. The remaining variables are interactions of the prior variables. Robust standard errors are used. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The results suggest that there is a conscious effect of the respondent's beliefs regarding the defendant's propensity to commit crime and the criminal justice system, but not the crime charged. The results also illustrate that these conscious effects fail to account for all of the difference in the presumptions of innocence assigned to black and white defendants.

First, the results in Table 12 indicate that there is no significant impact of the respondent's beliefs about crime charged, suggesting that respondents do not view arson and vandalism as crimes that are more or less likely to be committed by white individuals as compared to black individuals. However, the results tend toward showing that respondents view vandalism and arson as crimes committed more often by white individuals.

The Equation 4 results show that respondents who said the crime charged is important do not assign pre-evidence assessments of guilt that are significantly different from those assigned by respondents who said the crime charged is not important. Pre-evidence assessments of guilt also do not differ significantly between arson and vandalism cases. Equation 5 allows the effect of the importance of the crime charged to differ by the race of the defendant and the crime charged. It also allows the effect of the crime charged to differ based on the race of the defendant. The results illustrate that the effect of the importance of the crime charged differs significantly by the race of the defendant—finding the crime charged important increases the

pre-evidence assessment of guilt significantly more when the defendant is white. However, the effect of finding the crime charged important is insignificant for both white and black defendants. In other words, when the defendant is black, finding the crime charged important tends towards decreasing the pre-evidence assessment of guilt (but the effect is not significant), and when the defendant is white, finding the crime charged important tends towards increasing the pre-evidence assessment of guilt (but the effect is not significant). There are no significant effects related to the crime charged in the Equation 6 results, when the effect of defendant race on the impact of the importance of the crime charged is allowed to differ by the crime charged. Therefore, while there is a suggestion that respondents tend towards viewing arson and vandalism as crimes committed more often by white individuals, there is no significant impact of beliefs about the crime charged.

The Crime in the United States Report, published by the Federal Bureau of Investigation each year, contains information on the racial composition of arson and vandalism arrests, which can shed some light on the validity of respondents' beliefs. In 2017, there were a total of 842,976 adult arrests for property crime arson (Federal Bureau of Investigation 2017). Of these, 69.3 percent were of white individuals and 27.6 percent were of black individuals (Federal Bureau of Investigation 2017). That same year, there were 117,599 adult arrests for vandalism (Federal Bureau of Investigation 2017). Of these, 68.2 percent were of white individuals and 28 percent were of black individuals (Federal Bureau of Investigation 2017). In 2017, white individuals made up 76.6 percent of the United States population, and black individuals accounted for 13.4 percent of the United States population (U.S. Census Bureau d). These numbers illustrate that black individuals actually made up a disproportionate amount of arrests for property crime arson and vandalism in 2017.

However, because black individuals account for a disproportionate number of all arrests, it is arguably more illustrative to compare to the racial composition of all arrests rather than the racial composition of the population. In 2017, 7,537,750 adult arrests were made (Federal Bureau of Investigation 2017). White individuals accounted for 69.5 percent of those arrests, and black individuals accounted for 26.6 percent (Federal Bureau of Investigation 2017). Comparing these numbers to those for property crime arson and vandalism arrests illustrates that these crimes are actually fairly representative in terms of the racial composition of arrests. Therefore, respondents may correctly believe that white and black defendants are equally likely to commit the crimes of arson and vandalism. It is important to note, however, that there is a clear difference between arrests and crimes committed. The racial composition of actual crimes committed could be different from that of arrests, as arrests are influenced by victim reporting and police behavior.

The results in Table 12 also illustrate that the respondent's beliefs regarding the defendant's propensity to commit crime have a significant impact on the presumption of innocence. The Equation 4 results illustrate that individuals who think the defendant's propensity to commit crime is important assign significantly higher pre-evidence assessments of guilt than those who do not think that the defendant's propensity to commit crime is important.

The Equation 5 results, which allow the importance of the defendant's perceived propensity to commit crime to have a different effect based on the race of the defendant, reveal a surprising interaction effect. If the defendant is black, respondents who say that the defendant's propensity to commit crime is important give a significantly higher pre-evidence assessment of guilt (by about 6 percentage points) than those who say the defendant's propensity to commit crime is not important. However, when the defendant is white, respondents who say that the

defendant's propensity to commit crime is important do not significantly differ in their pre-evidence assessments of guilt from respondents who say the defendant's propensity to commit crime is not important. This result is surprising, because it works *against* the overarching result—that white defendants are given higher pre-evidence assessments of guilt than black defendants. That is, after accounting for the effect of differences in perceived propensity to commit crime, the effect of defendant race on the presumption of innocence is even greater, suggesting that mock jurors believe black defendants have a higher propensity to commit crime than white defendants. No new variables related to the propensity to commit crime are introduced in Equation 6, and the results do not meaningfully change.

Finally, the results in Table 12 illustrate that respondents' beliefs about the criminal justice system contribute to the differences in the presumptions of innocence given to white and black defendants. The Equation 4 results illustrate that individuals who say that their beliefs about the criminal justice system are very or somewhat important assign significantly higher pre-evidence assessments of guilt than those who do not, and individuals who distrust the justice system assign significantly lower pre-evidence assessments of guilt than those who do not.

Equation 5 adds interaction effects to allow the importance of the respondent's beliefs about the criminal justice system to have a different effect based on the race of the defendant and whether the respondent distrusts the justice system. It also allows the effect of distrusting the justice system to differ based on the race of the defendant. The results illustrate that distrusting the justice system decreases pre-evidence assessments of guilt significantly more when the defendant is black.

Equation 6 is the most complete specification because it additionally allows the effect of distrusting the justice system to differ based on the interaction of the race of the defendant and

the importance of justice system beliefs. The results illustrate that white defendants receive significantly higher pre-evidence assessments of guilt than black defendants from respondents who distrust the justice system: Those who think their opinions about the criminal justice system are important predict that white defendants are about 20 percentage points more likely to be guilty than black defendants, and those who do not think their criminal justice system beliefs are important predict that white defendants are about 10 percentage points more likely to be guilty than black defendants. However, the pre-evidence assessments of guilt assigned by respondents who do not distrust the justice system—regardless of whether they think their beliefs about the criminal justice system are important in assessing guilt pre-trial—do not differ significantly by the race of the defendant.

This suggests that some mock jurors may exhibit a bias against white defendants in order to correct for racial biases that they believe occur at other steps in the criminal justice process, a form of race-based jury nullification. This could be related to recent increases in media attention surrounding violent confrontations between black suspects of nonviolent crimes and police officers²⁹ or mock jurors own experiences or the experiences of their friends or family.³⁰ As a result of experience or exposure to such media coverage, jurors may believe, before hearing any evidence, that a black defendant charged with a nonviolent crime is less likely to be guilty than a white defendant.

Even after controlling for the conscious impact of the respondent's beliefs about the crime charged, the defendant's propensity to commit crime, and the criminal justice system,

²⁹ For example, in the highly publicized death of Eric Garner, the police approached Mr. Garner, a black man, on suspicion that he was selling individual cigarettes from packs without tax stamps and put him in a chokehold for about 15 to 19 seconds, resulting in his death (Baker, Goodman, & Mueller 2015). Similarly, the highly publicized shooting of Walter Scott occurred after a police officer pulled Mr. Scott over for a non-functioning brake light (Swaine 2015).

³⁰ Research has found that a number of characteristics including race, the level of neighborhood crime, and the reported quality of police-citizen encounters are related to trust and confidence in the police (Weitzer & Tuch 2005).

there remains a significant effect of defendant race on the presumption of innocence. This could be due to the fact that respondents are unconsciously influenced by one of these three factors, or it could be the result of an additional, unexplored factor.

2. Defendant Race Does Not Affect the Burden of Proof

Pre-existing racial biases may also affect the interpretation of the burden of proof. Recall that in the second study mock jurors were asked how sure they would need to be about the defendant’s guilt in order to convict. Respondents’ numerical interpretations of the burden of proof are not significantly affected by the defendant’s race regardless of the crime charged.

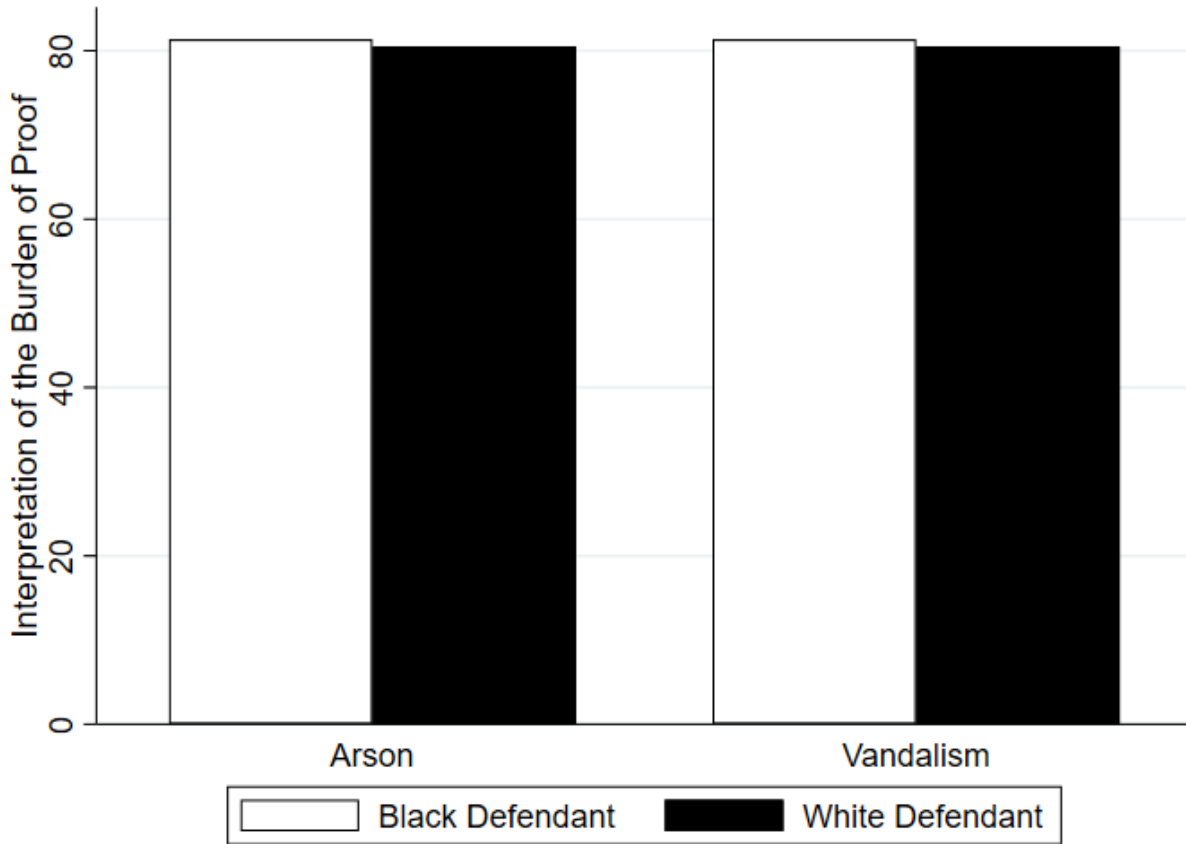
Table 13 presents and Figure 5 depicts the burden of proof results by defendant race.

Table 13. Burden of Proof Interpretations by Defendant Race

Case Scenario	All Defendants	White Defendants	Black Defendants
All Case Scenarios	80.95%	80.51%	81.37%
Vandalism	80.53%	81.12%	79.92%
Arson	81.36%	80.10%	81.61%

Notes: Numbers represent how sure mock jurors would need to be of the defendant’s guilt, on a scale of 0 to 100, inclusive, in order to return a verdict of guilty. Stars indicate significant differences between the numerical interpretations of the burden of proof for black and white defendants. *p<0.1 **p<0.05 ***p<0.01.

Figure 5. Burden of Proof by Defendant Race



As illustrated in Figure 1, theory suggests that mock jurors’ conceptualizations of the burden of proof may be influenced by their beliefs about the costs of wrongful acquittals, their beliefs about the costs of wrongful convictions, and their beliefs about the importance of the burden of proof. In order to explore the importance of these factors, the following equations are estimated using linear regression models. Again, all of the case scenarios are combined for these analyses.

$$BOPInterpretation_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 FactorImportance_i + \beta_3 Arson_i + \epsilon_i \tag{7}$$

$$BOPInterpretation_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 FactorImportance_i + \beta_3 WhiteDefendant_i * FactorImportance_i + \beta_4 Arson_i + \epsilon_i \tag{8}$$

$BOPInterpretation_i$ is a variable equal to the respondent's numerical interpretation of the burden of proof. $FactorImportance_i$ is a vector that includes three dummy variables equal to one if the respondent said the cost of a wrongful conviction, the cost of a wrongful acquittal, and whether the prosecutor has met the burden of proof are very important or somewhat important on a five point Likert scale, respectively. $WhiteDefendant_i * FactorImportance_i$ is a vector that contains the interactions of $WhiteDefendant_i$ and $FactorImportance_i$. $Arson_i$ is a dummy variable equal to one when the crime charged is arson.

Of course, it is possible that respondents are not aware of how important these factors are to their decision; these regressions will only capture the conscious, reported effects of these various factors. Regression results are presented in Table 14.

Table 14. Regression Analyses of Defendant Race, Perceived Importance of a Wrongful Conviction, Perceived Importance of a Wrongful Acquittal, and Perceived Importance of the Burden of Proof on Numerical Interpretations of the Burden of Proof

	Equation 7	Equation 8
White Defendant	-0.54	0.35
Wrongful Conviction Cost Important	4.48***	7.22***
Wrongful Acquittal Cost Important	-1.57	-3.29**
Burden of Proof Important	20.53***	19.91***
White Defendant * Wrongful Conviction Cost Important	—	-5.49**
White Defendant * Wrongful Acquittal Cost Important	—	3.47*
White Defendant * Burden of Proof Important	—	1.26
Arson	0.79	0.75
Observations	2,512	2,512
R-Squared	0.154	0.156

Notes: Standard ordinary least squares regressions are used. The outcome variable is the respondent's numerical interpretation of the burden of proof (i.e., how certain they said they would need to be of the defendant's guilt in order to convict), which is restricted to between 0 and 100, inclusive. White defendant is a dummy variable equal to one when the defendant is white. Wrongful Conviction Cost Important is a dummy variable equal to one if the respondent said the cost of a wrongful conviction is very important or somewhat important on a five point Likert scale. Wrongful Acquittal Cost Important is a dummy variable equal to one if the respondent said the cost of a wrongful acquittal is very important or somewhat important on a five point Likert scale. Burden of Proof Important is a dummy variable equal to one if the respondent said that whether the prosecutor meets the burden of proof is very important or somewhat important on a five point Likert scale. Arson is a dummy variable equal to one if the crime charged is arson. The remaining variables are interactions of the prior variables. Robust standard errors are used. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The Equation 7 results illustrate that, overall, caring about the costs of a wrongful conviction is associated with a significantly higher numerical interpretation of the burden of proof (by about 4 percentage points) and believing that the burden of proof is important is also associated with a significantly higher numerical interpretation of the burden of proof (by about 21 percentage points).

The Equation 8 results illustrate that, even though there is not a significant difference in the numerical interpretation of the burden of proof based on the race of the defendant, the costs of wrongful acquittals and convictions correlate differently with the numerical interpretation of the burden of proof based on the race of the defendant. If the defendant is black, viewing the possibility of a wrongful conviction as important is associated with a significantly higher numerical interpretation of the burden of proof (by about 7 percentage points). If the defendant is white, however, viewing a wrongful conviction as important is not significantly associated with the numerical interpretation of the burden of proof. If the defendant is black, viewing a wrongful acquittal as important is associated with a significantly lower numerical interpretation of the burden of proof (by about 3 percentage points). If the defendant is white, however, there is again no significant relationship. Caring about whether the prosecutor has met the burden of proof is associated with a significantly higher numerical interpretation of the burden of proof, but the effect does not differ significantly based on the race of the defendant.

These results suggest that, in the case of nonviolent property crimes, decisionmakers' race-based concerns about wrongful convictions and acquittals may cancel each other out, resulting in no significant differences in interpretations of the burden of proof by the race of the defendant.

D. Putting It All Together

Together, the results from the two experimental vignette studies suggest that in the context of a trial for a nonviolent property crime, decisionmakers exhibit a bias against white defendants when applying the presumption of innocence and when interpreting the evidence presented at trial, but not when interpreting the burden of proof. The presumption of innocence results are at least partially explained by race-based jury nullification. Controlling for evidence credibility and race-based weighing of evidence credibility eliminates the significant effect of race on conviction rates. This Part examines whether additionally controlling for the presumption of innocence and the burden of proof actually results in black defendants being worse off than white defendants or whether the impact of pre-existing racial biases overlaps with the impact of case-specific biases. In order to do this, the following equation is estimated using a linear probability model and the data from the first study, once for each criminal case scenario.

$$\begin{aligned}
 Guilty_i = & \alpha + \beta_1 WhiteDefendant_i + \beta_2 CredibilityRatings_i \\
 & + \beta_3 WhiteDefendant_i * CredibilityRatings_i \\
 & + \beta_4 EstPre-EvidenceGuilt_i + \beta_5 WhiteDefendant_i \\
 & * EstPre-EvidenceGuilt_i + \beta_5 EstBOPInterpretation_i \\
 & + \beta_6 WhiteDefendant_i * EstBOPInterpretation_i + \epsilon_i
 \end{aligned} \tag{9}$$

$Guilty_i$, $WhiteDefendant_i$, $CredibilityRatings_i$, and $WhiteDefendant_i *$

$CredibilityRatings_i$ are defined as in Equation 3. $EstPresumptionofInnocence_i$ is the pre-evidence guilt assessment that the respondent in the first study would likely have chosen, based on their demographic characteristics and the responses in the second study.³¹

³¹ Specifically, the predicted pre-evidence assessments of guilt were calculated by estimating an ordinary least squares regression of the following form using the data from the second study: $Pre-EvidenceGuilt_i = \alpha + \beta Demographics_i + \epsilon$, where $Pre-EvidenceGuilt$ is a variable equal to the pre-evidence guilt assessment and $Demographics_i$ is a vector that includes dummy variables for being female; being married; having at least a bachelor's degree; being employed; living the Northeast, Midwest, or West; being a democrat or republican; attending religious services at least once a month; identifying as black, Asian, another non-white race, or Hispanic; living in an urban area; having been the victim of a crime; having served on a jury; having been arrested or convicted of a felony or misdemeanor; and strongly or somewhat disagreeing that the justice system in the United

$EstBOPInterpretation_i$ is the numeric interpretation of the burden of proof that the respondent in the first study would likely have chosen, based on their demographic characteristics and the responses in the second study.³² $WhiteDefendant_i * EstPresumptionofInnocence_i$ is the interaction of $WhiteDefendant_i$ and $EstPresumptionofInnocence_i$, and $WhiteDefendant_i * EstBOPInterpretation_i$ is the interaction between $WhiteDefendant_i$ and $EstBOPInterpretation_i$. The results are presented in Table 15.

Table 15. Regression Analyses of Defendant Race, Evidence Believability, Pre-Evidence Guilt Assessments, and Burden of Proof Interpretations on Conviction Rates

	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
White Defendant	-1.32	0.74	2.65	-0.51
Prosecutor’s Version of Events	0.16***	0.12***	0.17***	0.15***
Defense’s Version of Events	-0.35***	-0.22***	-0.32***	-0.28***
Testimony of Eyewitness	0.14***	—	0.12***	—
Testimony of Insurance Agent	—	0.06**	—	0.03*
Testimony of Police Officer	0.05*	0.03	0.05	0.10***
Testimony of Defendant’s Friends	-0.21***	—	-0.29***	—
Testimony of Defendant’s Wife	—	-0.08*	—	-0.06
Testimony of Defendant	—	-0.19***	—	-0.21***
White Defendant * Prosecutor’s Version of Events	-0.09**	0.02	0.04	0.04
White Defendant * Defense’s Version of Events	0.02	-0.06	-0.01	0.01
White Defendant * Testimony of Eyewitness	-0.02	—	0.01	—
White Defendant * Testimony of Insurance Agent	—	-0.07*	—	0.02
White Defendant * Testimony of Police Officer	0.07*	0.05	0.04	-0.04

States works well. The same demographic characteristics were then plugged into the resulting estimated equation to predict pre-evidence assessments of guilt for respondents in the first study.

³² Specifically, the predicted burden of proof interpretations were calculated by estimating an ordinary least squares regression of the following form using data from the second study: $BOPInterpretation_i = \alpha + \beta Demographics_i + \epsilon$, where $BOPInterpretation_i$ is a variable equal to the numerical interpretation of the burden of proof and $Demographics_i$ is a vector that includes dummy variables for being female; being married; having at least a bachelor’s degree; being employed; living the Northeast, Midwest, or West; being a democrat or republican; attending religious services at least once a month; identifying as black, Asian, another non-white race, or Hispanic; living in an urban area; having been the victim of a crime; having served on a jury; having been arrested or convicted of a felony or misdemeanor; and strongly or somewhat disagreeing that the justice system in the United States works well. The same demographic characteristics were then plugged into the resulting estimated equation to predict burden of proof interpretations for respondents in the first study.

White Defendant * Testimony of Defendant's Friends	-0.06*	—	0.04	—
White Defendant * Testimony of Defendant's Wife	—	-0.02	—	-0.09
White Defendant * Testimony of Defendant	—	0.06	—	0.06
Estimated Pre-Evidence Guilt Assessment	-0.01	0.02	0.04**	0.04**
White Defendant * Estimated Pre-Evidence Guilt Assessment	0.002	0.001	-0.03	0.0003
Estimated Burden of Proof	-0.04***	-0.02	-0.01	-0.01
White Defendant * Estimated Burden of Proof	0.02	-0.01	-0.02	0.01
Observations	3,653	3,569	3,569	3,653
R-Squared	0.342	0.297	0.375	0.332

Notes: Standard ordinary least squares regressions are used. The outcome variable is a dummy variable equal to one if the defendant is found guilty. White defendant is a dummy variable equal to one if the defendant is white. Prosecutor's Version of Events, Defense's Version of Events, and all of the variables beginning with Testimony are dummy variables equal to one if the participant rated that aspect of trial very believable or somewhat believable on a five point Likert scale. Estimated Pre-Evidence Guilt Assessment is a predicted value of the respondent's pre-evidence assessment of the defendant's guilt. Estimated Burden of Proof is a predicted value of the respondent's numerical interpretation of the burden of proof. For more information on how these predicted values are calculated, see footnotes 31 and 32. The other variables are interactions. Standard errors are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

The results in Table 15 illustrate that the effects of differences in the presumption of innocence and the effects of differences in the credibility and weighing of the evidence are overlapping. After controlling for both, there is no significant effect of defendant race on conviction rates in any of the four criminal case scenarios.

Overall, the conviction results illustrate that in the context of trials for nonviolent property crimes, white defendants are between 4 and 10 percentage points more likely to be convicted. This bias is present among all racial and ethnic subgroups of mock jurors, but it is significantly larger among minority mock jurors. There is also evidence that the bias is not concentrated solely among mock jurors with characteristics distinct from actual judges and jurors.

Race-based jury nullification explains at least some of the difference in the conviction rates of black and white defendants. Defendant race significantly effects perceptions of the

credibility of the evidence and pre-evidence assessments of the defendant's guilt. Specifically, mock jurors who distrust the justice system assign a higher presumption of innocence to black defendants as compared to white defendants. The next Section will explore the effect of defendant race on sentencing.

VI. White Defendants Are Sentenced More Harshly Than Black Defendants When Mock Jurors Are Pooled

After rendering a verdict, all mock jurors in the first study were asked to assume the role of a sentencing judge, assume a jury had found the defendant guilty, and select a nonzero sentence for the defendant, regardless of their verdict decision. In vandalism cases, mock jurors rendered a sentence between one and five years, inclusive, and in arson cases, mock jurors rendered a sentence between one and ten years, inclusive. This Section will examine the impact of defendant race on assigned sentence length.

The motive scenarios result in greater sentences than the eyewitness scenarios, and, unsurprisingly, arson cases result in higher sentences than vandalism cases.³³ Figures 6 and 7 display the distribution of sentences in each of the four case scenarios. In all of the case scenarios, the distribution is left-skewed, indicating a preference for lighter sentences. In the eyewitness vandalism case scenario, around 70 percent of respondents assign the minimum sentence of one year. In the motive vandalism case scenario, just over 60 percent of respondents assign the minimum sentence of one year. Among mock jurors who found the defendant guilty, just over 60 percent return the minimum sentence in the eyewitness vandalism case scenario, and about 45 percent return the minimum sentence in the motive vandalism case scenario.

³³ Some of this difference could be due to an anchoring effect because the maximum sentence was higher for arson than for vandalism (Enough & Mussweiler 2001). Such an anchoring effect is not unique to the experimental vignette study context, however, as judges are also exposed to minimum and maximum sentences imposed by Congress and sentence recommendations presented by the defense and prosecution.

There is significantly more variation in the sentences assigned in arson cases. In the eyewitness arson case scenario, just under 25 percent of respondents assign the minimum sentence of a year, and the median sentence is two years. In the motive arson case scenario, around 16 percent of respondents assign the minimum sentence of a year, and the median sentence is 2.9 years. Of the jurors that returned a verdict of guilty, about 14 percent assign the minimum sentence in the eyewitness arson case scenario, and about 5 percent assign the minimum sentence in the motive arson case scenario. About 68 percent of participants feel somewhat confident or very confident in their sentencing decision.

Figure 6. Distribution of Sentences Assigned in Vandalism Cases

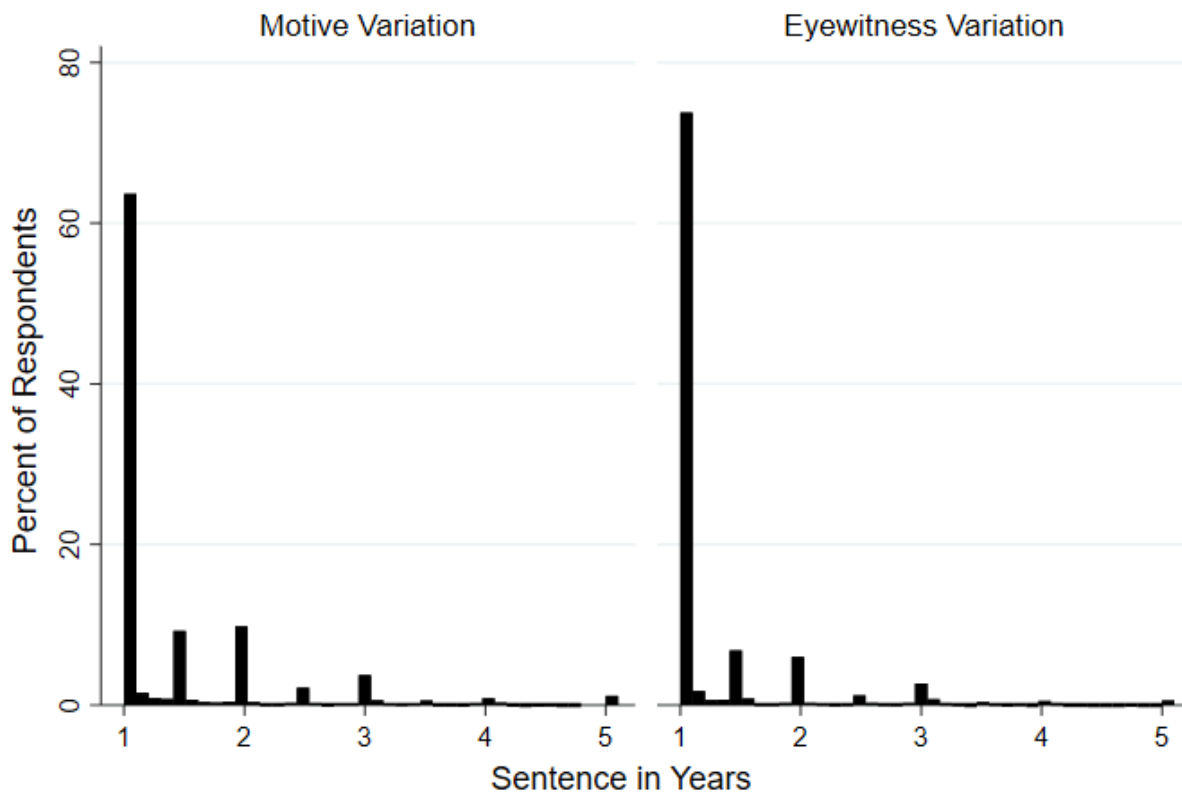
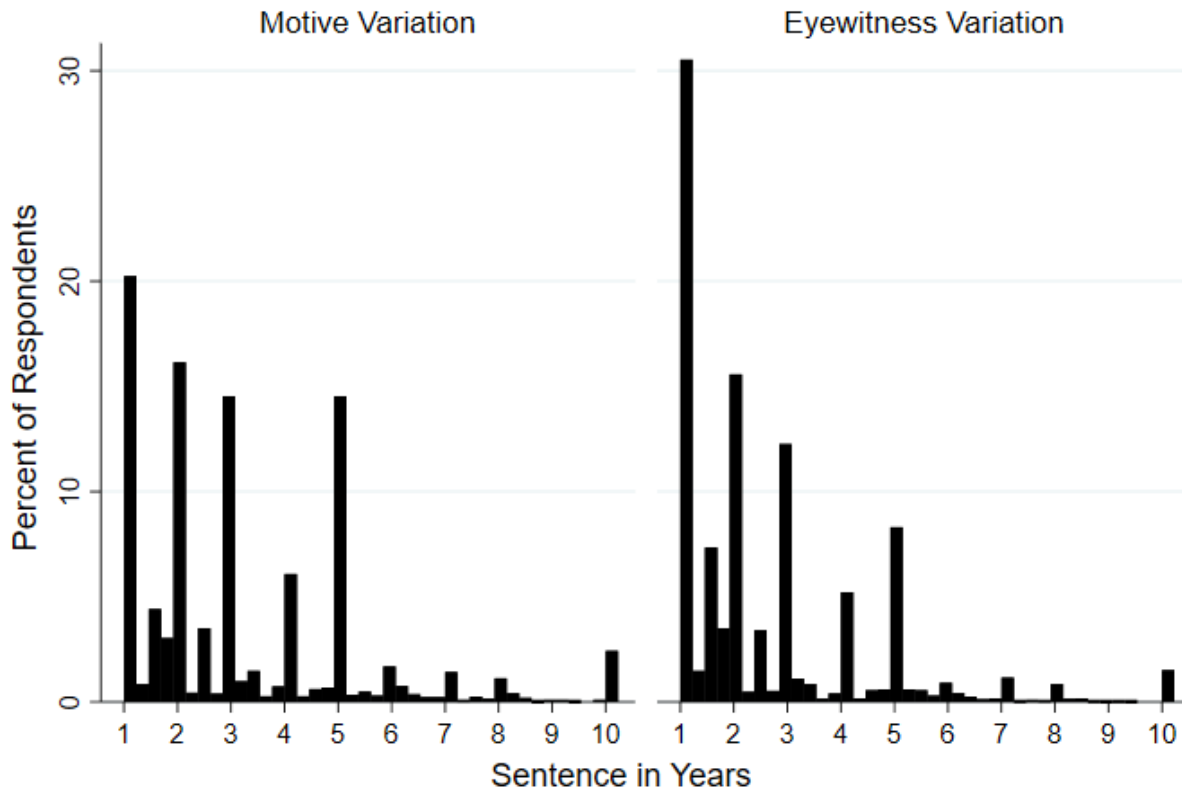


Figure 7. Distribution of Sentences Assigned in Arson Cases



Despite the large number of mock jurors who return the minimum sentence, there is a significant effect of defendant race on sentence length in arson cases. In the motive arson case scenario, white defendants are sentenced to an average of 95 more days in prison than black defendants. In the eyewitness arson case scenario, white defendants are sentenced to an average of 47 more days in prison than black defendants. Table 16 summarizes sentencing decisions by defendant race for each of the case scenarios in the first study.

Table 16. Sentence Length by Defendant Race

Case Scenario	All Defendants	White Defendants	Black Defendants
All Case Scenarios	2.14	2.19***	2.10
Eyewitness Vandalism	1.31	1.30	1.32
Motive Vandalism	1.45	1.47	1.43
Eyewitness Arson	2.62	2.69**	2.56
Motive Arson	3.19	3.32***	3.06

Notes: In vandalism cases, participants were restricted to sentences between 1 and 5 years, inclusive; in arson cases, participants were restricted to sentences between 1 and 10 years, inclusive. Stars indicate significant differences between the average sentences assigned to black and white defendants. *p<0.1 **p<0.05 ***p<0.01.

A. Mock Juror Heterogeneity: White Defendants Are Sentenced More Harshly Than Black Defendants by Mock Jurors Who Acquitted and Black Defendants Are Sometimes Sentenced More Harshly Than White Defendants by Mock Jurors Who Convicted

The effect of defendant race on verdict determinations differs significantly by the race of the mock juror. In order to explore whether the effect of race on sentencing decisions is similarly heterogeneous, Table 17 contains the differences in average sentences assigned to white and black defendants by mock juror race, ethnicity, and southern residency. Only white mock jurors exhibit a significant racial bias in the eyewitness vandalism case scenario, and this bias, though only marginally significant, is against black defendants rather than white defendants. Black, Hispanic, and southern mock jurors exhibit a significant bias against white defendants in the motive vandalism case scenario. White, non-Hispanic, and non-southern mock jurors exhibit a bias against white defendants in the eyewitness arson case scenario. Finally, white, Asian, non-Hispanic, southern, and non-southern mock jurors all exhibit a significant bias against white defendants in the motive arson case scenario.

Table 17. Disparities in Sentences Given to White and Black Defendants by Mock Juror Race, Ethnicity, and Residence in the South Census Region

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson	Observations
White Only	0.08***	-0.04*	0.02	0.15**	0.22***	12,630
Black Only	0.18*	0.06	0.19**	0.26	0.35	1,396
Asian Only	0.13	0.04	0.004	0.13	0.53**	1,012
Hispanic	0.02	-0.06	0.23**	-0.29	0.15	1,268
Non-Hispanic	0.11***	-0.02	0.03	0.19***	0.26***	14,472
Southern	0.15***	-0.01	0.09**	0.15	0.36***	6,220
Not Southern	0.07*	-0.02	0.01	0.13*	0.20**	9,920

Notes: Numbers represent the average sentence for white defendants minus the average sentence for black defendants assigned by respondents of the given race, ethnicity, or Census region in the given case scenario. Only individuals who report a single race are included in this analysis. Individuals are considered southern if their state of residence is in the south Census region and not southern otherwise. *p<0.1 **p<0.05 ***p<0.01.

In order to verify whether racial sentencing disparities differ significantly by the race, ethnicity, or southern residency of the mock juror, Equation 10 is estimated for the subgroups of mock jurors listed in Table 17.

$$Sentence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \epsilon_i \quad (10)$$

$Sentence_i$ is equal to the sentence assigned by the mock juror, $WhiteDefendant_i$ is a dummy variable equal to one when the defendant is white, $Arson_i$ is a dummy variable equal to one when the crime charged is arson, and $Eyewitness_i$ is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering around a motive. In order to maintain adequate sample sizes, all four case scenarios are pooled for this analysis. Standard errors are clustered at the individual level.

Once these regressions are estimated, the β_1 coefficients for related groups of mock jurors are tested for equivalency. Unlike conviction rates, sentence disparities do not differ significantly by the race of the mock juror. They also do not differ significantly by mock juror ethnicity or southern residency.³⁴

It is also important to determine whether the bias against white defendants in sentencing is concentrated among individuals with characteristics distinct from actual judges and jurors. Table 18 reports the differences in average sentences given to white and black defendants by those with and without the education level and approximate income level of judges, those with and without the characteristics most common among seated jurors, and those who have and have not previously served on a jury. As before, based on data from the Bureau of Labor Statistics' Occupational Employment Statistics survey, individuals are considered to have the education

³⁴ Even though the magnitude of the race coefficient is not significantly different based on the race, ethnicity, or southern residency of the mock juror, likelihood ratio tests indicate that including mock juror race and southern residency (but not mock juror ethnicity) as explanatory variables in Equation 10 does significantly increase the fit of the model.

level and approximate income level of judges if they have a doctoral or professional degree and an income above \$75,000 per year (Bureau of Labor Statistics 2018). Based on data from a survey administered to seated jurors in four courts by Hannaford-Agor et al. (2003) and the demographic characteristics of participants who report prior jury service in this study (presented in Table 6), individuals are considered to have the most common characteristics of seated jurors if they have at least some college education and are employed.

Table 18. Disparities in Sentences Given to White and Black Defendants by Judicial and Juror Demographics and Prior Jury Service

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson	Observations
Prior Jury Service	0.06	-0.001	0.13**	-0.15	0.33*	2,200
No Prior Jury Service	0.10***	-0.02	0.02	0.19***	0.25***	13,940
Characteristics Similar to Seated Jurors	0.09***	-0.02	0.03	0.11	0.31***	12,260
Characteristics Distinct from Seated Jurors	0.11**	0.004	0.07	0.24**	0.09	3,880
Education and Income Similar to a Judge	0.15	0.02	0.12	0.25	0.39	1,186
Education or Income Distinct from a Judge	0.09***	-0.02	0.03	0.13**	0.25***	14,880

Notes: Numbers represent the average sentence for white defendants minus the average sentence for black defendants assigned by respondents with the listed characteristic in the given case scenario. Individuals are considered to have education and income similar to a judge if they have a doctoral or professional degree and if their income is \$75,000 per year or more. Individuals are considered to have characteristics similar to seated jurors if they have at least some college education and are employed. Mock jurors who elect not to report their income are not included in either the group of mock jurors with education and income similar to judges or the group of mock jurors with education and income distinct from judges. *p<0.1 **p<0.05 ***p<0.01.

Those with income and education levels similar to judges do not assign significantly different sentences based on the race of the defendant. However, as before, the lack of significance could be related to reduced sample sizes. In fact, the raw disparities in the average sentences assigned to white and black defendants are actually larger in magnitude among those

with education and income levels similar to judges. In the motive variations, those who have served on a jury exhibit a greater bias against white defendants than those who have not, possibly indicating that motive is a more salient issue to prior jurors; in the eyewitness vandalism case scenario, the opposite is true. Those with characteristics similar to seated jurors exhibit a greater bias against white defendants in the motive arson case scenario but exhibit a smaller bias against white defendants in the eyewitness arson case scenario. These results do not provide strong evidence that the bias against white defendants in terms of sentencing is concentrated only among mock jurors who are dissimilar from judges and actual jurors.

Finally, it is important to explore whether the effects of defendant race on sentencing differ based on the verdict rendered. While it may seem strange to ask mock jurors who acquitted the defendant to assign a nonzero prison sentence, this is a situation that many real-world judges may face. In a case where a defendant is found guilty by a jury, the judge might think that, had the trial been a bench trial, he or she would have acquitted. Nevertheless, if the verdict is reasonable, the judge will need to accept the verdict and assign a sentence.

It is possible, however, that judges assign sentences differently in cases where they believe that the defendant is guilty and cases where they would have acquitted the defendant.³⁵ In the context of this study, this may translate into defendant race having a different effect on the assigned sentence depending on mock jurors' beliefs about the defendant's guilt, especially given that the burden of proof results indicate mock jurors are more worried about both wrongful convictions and wrongful acquittals when the defendant is black.

³⁵ This idea is related to the residual doubt doctrine, which in some cases may allow juries to consider residual doubt about the defendant's guilt in determining whether to impose the death penalty, even after they have found the defendant guilty beyond a reasonable doubt (Fisher 2012; Koosed 2001).

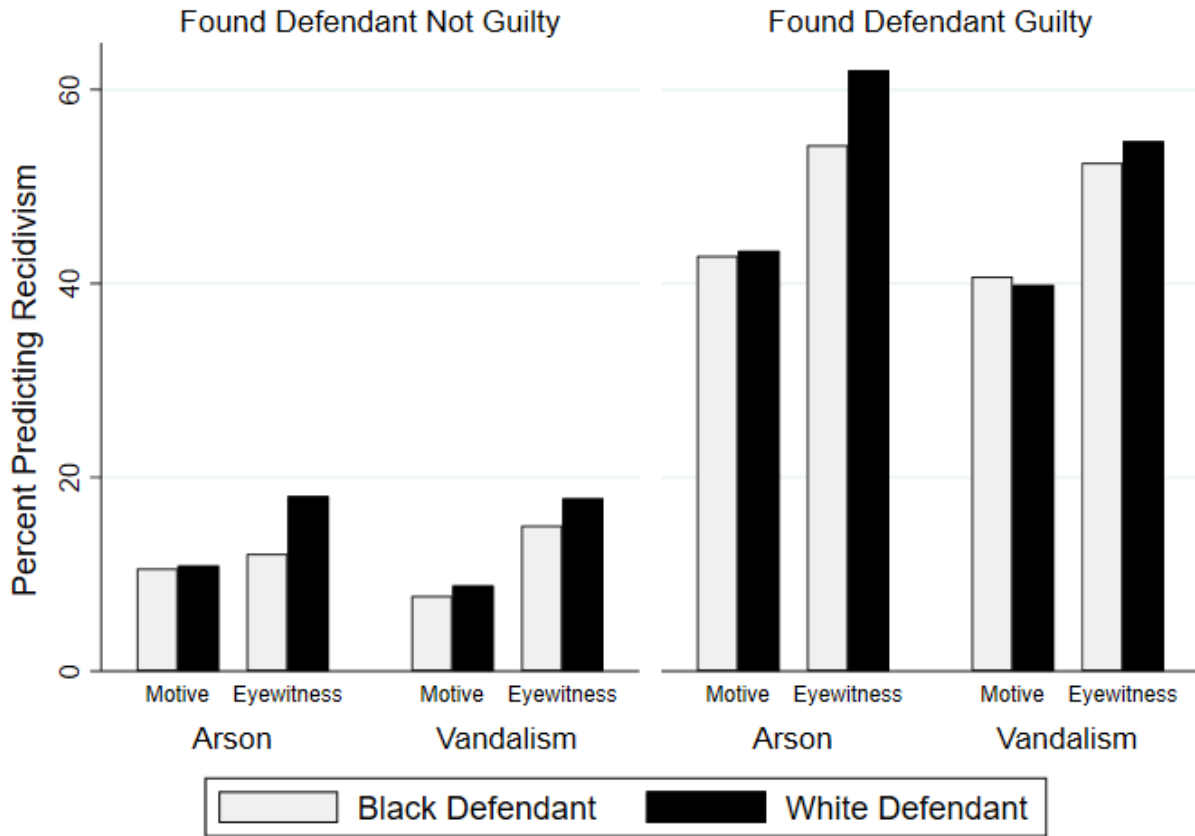
In fact, in all of the case scenarios except eyewitness vandalism and when all of the case scenarios are combined, mock jurors who found the defendant not guilty sentence white defendants significantly more harshly than black defendants (by 15 days in motive vandalism cases, 84 days in eyewitness arson cases, 99 days in motive arson cases, and 37 days when all of the case scenarios are combined). Conversely, in the eyewitness variations, mock jurors who found the defendant guilty sentence black defendants significantly more harshly than white defendants (by 33 days in vandalism cases and 58 days in arson cases). Table 19 contains and Figure 8 depicts the racial sentencing disparities by verdict rendered. The next Part will examine one mechanism through which race may affect sentence length—recidivism predictions.

Table 19. Disparities in Sentences Given to White and Black Defendants by Verdict Rendered

	Found Defendant Guilty	Found Defendant Not Guilty
All Case Scenarios	-0.06	0.10***
Eyewitness Vandalism	-0.09*	0.01
Motive Vandalism	-0.04	0.04*
Eyewitness Arson	-0.16*	0.23***
Motive Arson	-0.04	0.27***
Observations	5,782	10,358

Notes: Numbers represent the average sentence for white defendants minus the average sentence for black defendants assigned by respondents who returned the particular verdict. *p<0.1 **p<0.05 ***p<0.01.

Figure 8. Sentence Length by Defendant Race and Verdict Rendered



B. Race-Based Differences in Perceptions About Recidivism Account for Sentencing Disparities Among Mock Jurors Who Convicted but Not Among Mock Jurors Who Acquitted

The sentencing results illustrate that white defendants receive significantly longer prison sentences than black defendants from mock jurors who acquitted, and black defendants sometimes receive significantly longer prison sentences than white defendants from mock jurors who convicted. Because decisionmakers are typically given more discretion when assigning a sentence than when reaching a verdict, identifying the mechanism through which race is impacting sentencing requires sifting through many possible explanations. One such explanation

is recidivism predictions. Decisionmakers might believe that defendants of a particular race are more likely to reoffend and sentence them to longer prison terms as a result.

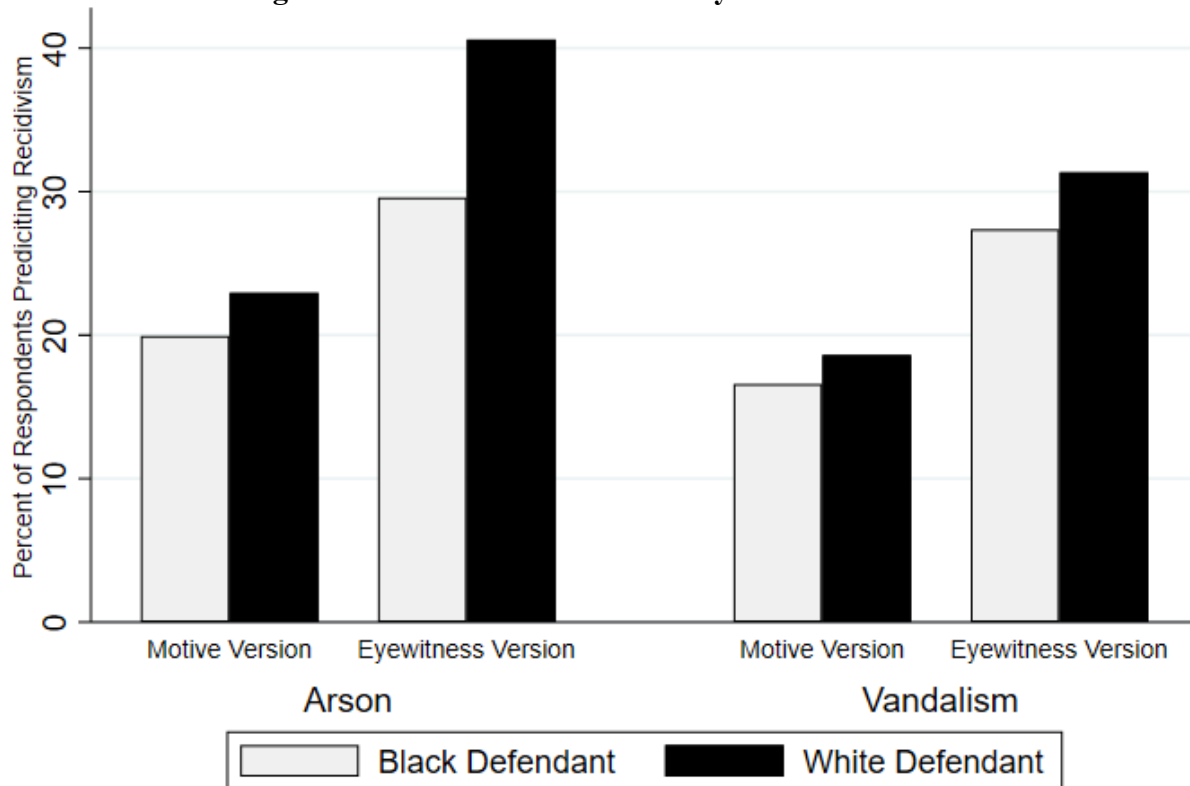
Recall that after sentencing the defendant, participants in the first study were asked to rate how likely they thought it was that the defendant would commit a crime in the next five years if acquitted. In all of the case scenarios—even the vandalism cases where there was not a significant difference in the sentences assigned to black and white defendants—and when all of the case scenarios are combined, participants predict that white defendants would commit a crime significantly more than they predict black defendants would commit a crime. The largest disparity occurs in the eyewitness arson case scenario, where about 41 percent of participants predict that white defendants will commit a crime in the next five years if acquitted and only about 30 percent of participants predict that black defendants will commit a crime in the next five years if acquitted. The smallest disparity occurs in the motive vandalism case scenario where about 19 percent of participants predict that white defendants will commit a crime in the next five years if acquitted and about 17 percent of participants predict that black defendants will commit a crime in the next five years if acquitted. Table 20 summarizes and Figure 9 depicts these predictions by defendant race.

Table 20. Recidivism Predictions by Defendant Race

Case Scenario	All Defendants	White Defendants	Black Defendants
All Case Scenarios	25.86%	28.41%***	23.33%
Eyewitness Vandalism	29.43%	31.35%***	27.39%
Motive Vandalism	17.57%	18.61%*	16.59%
Eyewitness Arson	35.00%	40.60%***	29.59%
Motive Arson	21.44%	22.94%**	19.95%

Notes: Stars indicate significant differences in the percent of respondents who predict that the defendant is very likely or somewhat likely to recidivate in the next five years if acquitted on a five point Likert scale for black and white defendants. *p<0.1 **p<0.05 ***p<0.01.

Figure 9. Recidivism Predictions by Defendant Race



While the effect of defendant race on sentencing differs based on the verdict rendered, the effect of defendant race on recidivism predictions does not. Among mock jurors who found the defendant guilty and mock jurors who found the defendant not guilty, white defendants are seen as significantly more likely to reoffend, particularly in the eyewitness arson case scenario. Table 21 contains the racial disparities in recidivism predictions by the verdict rendered.

Table 21. Racial Disparities in the Percent of Mock Jurors Predicting That the Defendant Will Recidivate by Verdict Rendered

	Found Defendant Guilty	Found Defendant Not Guilty
All Case Scenarios	0.03**	0.02***
Eyewitness Vandalism	0.02	0.03*
Motive Vandalism	-0.01	0.01
Eyewitness Arson	0.08***	0.06***
Motive Arson	0.01	0.003
N	5,782	10,358

Notes: Numbers represent the difference in the percent of respondents who predict that white defendants are very or somewhat likely to reoffend in the next five years if acquitted on a five point Likert scale and the percent of

These results indicate that mock jurors who found the defendant guilty sentence black defendants more harshly despite viewing white defendants as more likely to reoffend. They also indicate that mock jurors who found the defendant not guilty may sentence white defendants more harshly based on a perception that white defendant are more likely to reoffend than black defendants. In order to determine how much the racial disparities in sentencing are accounted for by the racial disparities in recidivism predictions, the following equations are estimated using linear regression models, once for mock jurors who found the defendant guilty and once for mock jurors who found the defendant not guilty.

$$Sentence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 PercievedRecidivist_i + \beta_3 Arson_i + \beta_4 Eyewitness_i + \epsilon_i \quad (11)$$

$$Sentence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 PercievedRecidivist_i + \beta_3 WhiteDefendant_i * PercievedRecidivist_i + \beta_4 Arson_i + \beta_5 Eyewitness_i + \epsilon_i \quad (12)$$

$Sentence_i$ is equal to the sentence assigned by the mock juror and $WhiteDefendant_i$ is a dummy variable equal to one when the defendant is white. $PercievedRecidivist_i$ is a dummy variable equal to one if the respondent predicted that the defendant is very likely or somewhat likely to commit a crime in the next five years if acquitted. $WhiteDefendant_i * PercievedRecidivist_i$ is the interaction of $WhiteDefendant_i$ and $PercievedRecidivist_i$. $Arson_i$ is a dummy variable equal to one when the crime charged is arson, and $Eyewitness_i$ is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering around a motive. In order to maintain adequate sample sizes, all four case scenarios are pooled for this analysis. Standard errors are clustered at the individual level.

Table 22 contains the regression results for Equation 11, and Table 23 contains the regression results for Equation 12.

Table 22. Regression Analyses of Defendant Race and Recidivism Predictions on Sentence Length by Verdict Rendered

	Found Defendant Guilty	Found Defendant Not Guilty
White Defendant	-0.11**	0.13***
Perceived Recidivist	0.64***	0.62***
Arson	1.82***	1.28***
Eyewitness	-0.68***	-0.38***
Observations	5,782	10,358
R-Squared	0.260	0.236

Notes: Standard ordinary least squares regressions are used. The outcome variable is the sentence assigned, which is restricted to between one and five years, inclusive, for vandalism cases and between one and ten years, inclusive, for arson cases. White defendant is a dummy variable equal to one when the defendant is white. Perceived recidivist is a dummy variable equal to one if the participant predicted that the defendant is very likely or somewhat likely to commit a crime in the next five years if acquitted on a five point Likert scale. Arson is a dummy variable equal to one when the crime charged is arson, and eyewitness is a dummy variable equal to one when the case revolves around the testimony of an eyewitness rather than a motive. Standard errors are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

Table 23. Regression Analyses of Defendant Race, Recidivism Predictions, and Race-Based Weighing of Recidivism Predictions on Sentence Length by Verdict Rendered

	Found Defendant Guilty	Found Defendant Not Guilty
White Defendant	0.04	0.14***
Perceived Recidivist	0.81***	0.67***
White Defendant * Perceived Recidivist	-0.30***	-0.11
Arson	1.82***	1.28***
Eyewitness	-0.67***	-0.38***
Observations	5,782	10,358
R-Squared	0.262	0.236

Notes: Standard ordinary least squares regressions are used. The outcome variable is the sentence assigned, which is restricted to between one and five years, inclusive, for vandalism cases and between one and ten years, inclusive, for arson cases. White defendant is a dummy variable equal to one when the defendant is white. Perceived recidivist is a dummy variable equal to one if the participant predicted that the defendant is very likely or somewhat likely to commit a crime in the next five years if acquitted on a five point Likert scale. White Defendant * Perceived Recidivist is the interaction of these variables. Arson is a dummy variable equal to one when the crime charged is arson, and eyewitness is a dummy variable equal to one when the case revolves around the testimony of an eyewitness rather than a motive. Standard errors are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

Unsurprisingly, the results in the first column of Table 22 indicate that the bias against black defendants by mock jurors who found the defendant guilty becomes larger and more significant after controlling for recidivism predictions. Even though these mock jurors view

white defendants as more likely to recidivate, they sentence black defendants to longer prison terms. The results in the second column are more surprising. Controlling for recidivism predictions does not decrease the size or significance of the effect of defendant race among mock jurors who found the defendant not guilty; if anything, it increases the size of the effect. This means that mock jurors who found the defendant not guilty are sentencing white defendants significantly more harshly than black defendants for some reason other than perceiving them as more likely to recidivate.

The results in Table 23, which allow the effect of recidivism predictions on sentence length to differ based on the race of the defendant, provide some additional insight. The significant effect of defendant race on sentence length among mock jurors who found the defendant guilty is accounted for by differences in the weighing of recidivism predictions: These mock jurors believe that white defendants are more likely to reoffend, but they sentence black defendants to longer prison terms because they punish black defendants for a given perceived likelihood of recidivism significantly more harshly than white defendants. This is not the case among mock jurors who found the defendant not guilty. Those mock jurors weigh the likelihood of recidivism the same regardless of the race of the defendant.

If mock jurors who found the defendant not guilty are making sentencing decisions in order to rehabilitate the defendant, deter future crime, or obtain retribution, the residual bias against white defendants in sentencing could be based on a belief that white defendants are more blameworthy or a belief that white defendants need to be sentenced to a greater amount of time in prison to be properly rehabilitated or deterred. It is also possible that the remaining disparity in sentencing represents another form of race-based jury nullification. Participants who found black defendants not guilty may believe that those defendants were convicted based on their race. If

this is the case, they may sentence these black defendant to less time than they would an equally situated white defendant in order to correct for what they perceive to be a racial injustice. This study is not able tease out these potential sources of the residual bias, but it is important to note that whatever the source of the bias, it—as well as the bias exhibited by mock jurors who found the defendant guilty—is entirely race-based; the facts of the case and the background facts about the defendant in this experiment do not vary based on the race of the defendant.

VII. Ruling Out the Attractiveness Alternative

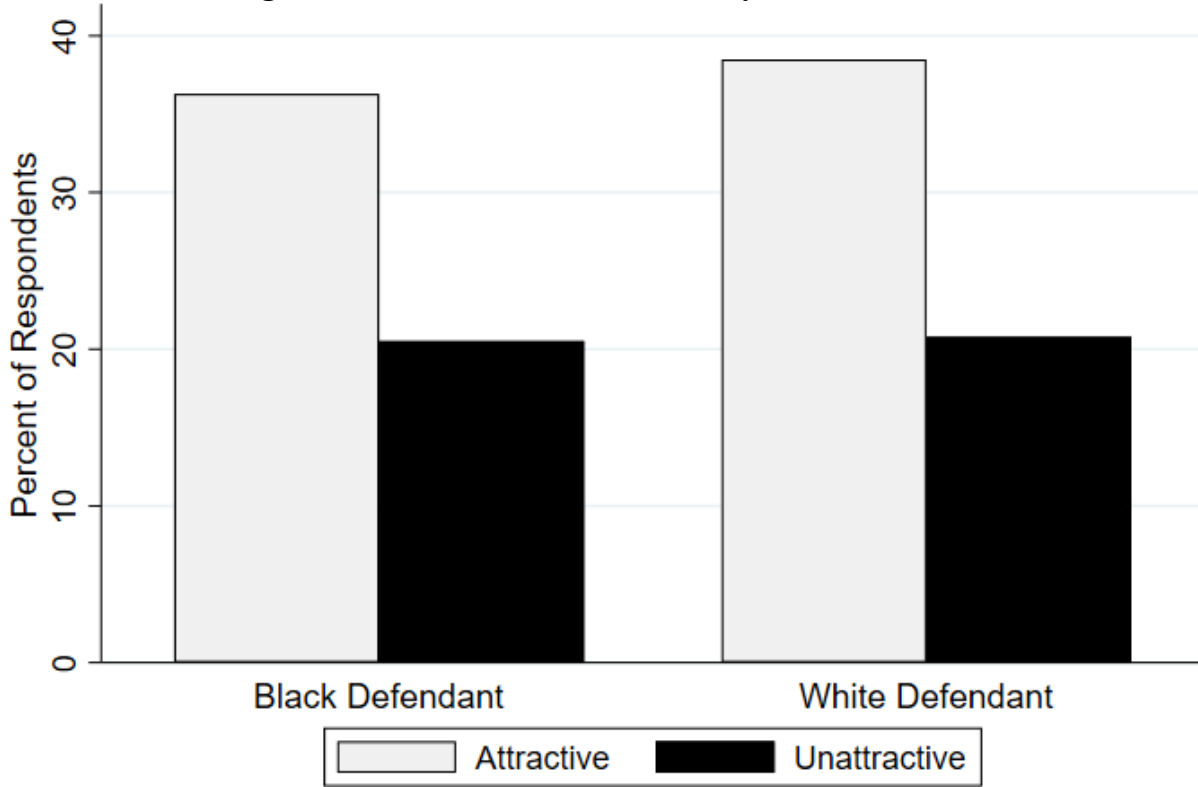
Recall that participants in the second study were also asked an attractiveness question, which was designed to test whether attractiveness could serve as alternative explanation for the effects of race discussed above. The percent of respondents who found defendants to be attractive and unattractive does not differ significantly by the race of the defendant. Table 24 presents and Figure 10 illustrates second study participants’ assessments of defendant attractiveness and unattractiveness by race. A respondent is coded as having found a defendant unattractive if he or she said the defendant was very or somewhat unattractive on a five point Likert scale and as having found a defendant attractive if he or she said the defendant was very or somewhat attractive on a five point Likert scale.

Table 24. Perceptions of Defendant Attractiveness by Defendant Race

	All Defendants	White Defendants	Black Defendants
Attractive	37.41%	38.50%	36.30%
Unattractive	20.65%	20.79%	20.51%

Notes: Respondents who find the defendant very or somewhat attractive are coded as finding the defendant attractive and respondents who find the defendant very or somewhat unattractive are coded as finding the defendant unattractive. The differences between white and black defendants are not significant.

Figure 10. Perceived Attractiveness by Defendant Race



In order to ensure that differences in defendant attractiveness do not account for the racial biases presented above, the following equations are estimated using linear probability and linear regression models.

$$Guilty_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \epsilon_i \quad (13)$$

$$Guilty_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \beta_4 Attractive_i + \epsilon_i \quad (14)$$

$$Guilty_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \beta_4 Unattractive_i + \epsilon_i \quad (15)$$

$$Sentence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \epsilon_i \quad (16)$$

$$Sentence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \beta_4 Attractive_i + \epsilon_i \quad (17)$$

$$Sentence_i = \alpha + \beta_1 WhiteDefendant_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \beta_4 Unattractive_i + \epsilon_i \quad (18)$$

The variables in Equation 13 are defined in the same manner as Equation 1, the variables in Equation 16 are defined in the same manner as Equation 10, *Attractive_i* is a variable equal to the percent of respondents in the second study that find the defendant somewhat or very attractive on a five point Likert scale, and *Unattractive_i* is a variable equal to the percent of respondents in the second study that find the defendant somewhat or very unattractive on a five point Likert scale.³⁶ The results are presented in Table 25.

Table 25. Regression Analyses of Defendant Race and Perceived Defendant Attractiveness on Case Outcomes

	Guilt	Guilt Controlling for Attractive	Guilt Controlling for Unattractive	Sentence	Sentence Controlling for Attractive	Sentence Controlling for Unattractive
White Defendant	0.07***	0.07***	0.07***	0.11***	0.10***	0.11***
Arson	0.08***	0.08***	0.08***	1.53***	1.53***	1.53***
Eyewitness	0.09***	0.09***	0.09***	-0.35***	-0.33***	-0.32***
Attractive	—	-0.0002	—	—	0.007*	—
Unattractive	—	—	0.0002	—	—	-0.01*
Observations	16,140	16,140	16,140	16,140	16,140	16,140
R-Squared	0.021	0.021	0.021	0.217	0.217	0.217

Notes: Standard ordinary least squares regressions are used. The outcome variable is either a dummy variable equal to one if the defendant is found guilty or the sentence assigned, which is restricted to between one and five years, inclusive, for vandalism cases and between one and ten years, inclusive, for arson cases. White defendant is a dummy variable equal to one when the defendant is white. Arson is a dummy variable equal to one when the crime charged is arson. Eyewitness is a dummy variable equal to one when the case revolves around eyewitness testimony rather than a motive. Attractive is a variable equal to the percent of respondents in the second study that find the defendant somewhat or very attractive on a five point Likert scale. Unattractive is a variable equal to the percent of respondents in the second study that find the defendant somewhat or very unattractive on a five point Likert scale. Standard errors are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

Neither attractiveness nor unattractiveness is significantly correlated with conviction decisions, and including controls for them does not materially change the magnitude or the significance of the effect of defendant race. Attractiveness and unattractiveness are marginally

³⁶ Controlling for attractiveness using the average scores assigned to pictures by a separate sample of respondents in this manner is consistent with other beauty literature (Biddle & Hamermesh 1998). This method assumes common standards of beauty in the population. Previous papers that have examined this assumption have found it to hold (see Hamermesh & Biddle 1994; Hatfield & Sprecher 1986; Biddle & Hamermesh 1998).

significantly correlated with sentencing decisions. However, including controls for them does not diminish the magnitude or significance of the effect of defendant race. Overall, these results confirm that race (rather than attractiveness) is the driving force behind the results presented above.

VIII. Discussion and Conclusion

The results in this Chapter illustrate that at trials for nonviolent property crimes, white defendants are convicted at significantly higher rates than black defendants. This discrepancy is at least partially accounted for by a form of race-based jury nullification—mock jurors’ beliefs about biases at other stages of the criminal justice system lead them to assign a stronger presumption of innocence to black defendants and view the evidence presented against black defendants more cautiously. Surprisingly, given the overarching result that white defendants are convicted more often than black defendants, mock jurors perceive that black defendants have a higher propensity to commit crime generally—the racial bias against white defendants is exaggerated by controlling for the importance of perceptions about the defendant’s propensity to commit crime.

The results also illustrate that white defendants are sentenced significantly more harshly than black defendants by mock jurors who found the defendant not guilty, and black defendants are sometimes sentenced significantly more harshly than white defendants by mock jurors who found the defendant guilty. Differential weighing of recidivism predictions explains the bias against black defendants exhibited by mock jurors who rendered a guilty verdict: Even though these mock jurors view white defendants as more likely to reoffend, they punish black defendants significantly more for a given perception of recidivism likelihood. Conversely, the bias against white defendants exhibited by mock jurors who found the defendant not guilty

remains large and significant after controlling for recidivism predictions. This residual bias could be the result of race-based differences in perceived culpability or capacity for rehabilitation. Alternatively, it could be the result of more race-based jury nullification if mock jurors who found black defendants not guilty assume that those defendants could only have been convicted based on racism and sentence them particularly leniently to correct for this perceived injustice.

In light of the fact that black individuals, particularly young black men, are disproportionately represented at every stage of the criminal justice system, these results are surprising. However, the apparent discrepancy between nationwide statistics and the results of this study can be explained in a number of ways. First, other biases in the system favoring white individuals may more than counteract the effect of any conviction or sentence disparities disfavoring them. For example, arrests and investigations by police, charging and plea decisions by prosecutors, and judicial decisions including setting bail, determining probable cause, and ruling on the admissibility of evidence could be influenced by defendant race. The apparent discrepancy could also result from the fact that racial disparities in conviction rates and sentencing in violent crime cases run in the opposite direction. It is also possible that black defendants commit more nonviolent crimes than white defendants (although arrest statistics seem to indicate that at least arson and vandalism are fairly race-neutral crimes). Finally, the apparent discrepancy could be based on the fact that criminal justice outcomes are heavily influenced by variables other than race that are not included in this study and that happen to be highly correlated with race, such as socioeconomic status.

One potential limitation to consider is the external validity of the results. The mock jurors did not witness a real trial with all of its nuances or a real in-person defendant, nor did they deliberate with others. The existing evidence on the impact of jury deliberations on biases is

mixed. Bernard (1979) found white jurors exhibiting bias towards black defendants to be uncompromising during deliberations and Kramer, Kerr, and Carroll (1990) found that the biasing effect of emotional publicity increased following jury deliberations. Conversely, Sommers (2006) found that deliberations reduce bias against black defendants when the jury is racially mixed and others (Kerwin & Shaffer 1994, London & Nunez 2000, Ruva et al. 2007) have found that the biasing effect of inadmissible information is reduced following deliberations. The best evidence that pre-deliberation decisions made by mock jurors are representative of post-deliberation jury decisions comes from a series of mock jury studies by Davis et al. (1975) that compared actual pre-deliberation verdict decisions in criminal cases with the jury verdicts reached after deliberations and found that the best way to transform pre-deliberation verdicts to post-deliberation verdicts is to predict that any verdict favored by two-thirds of the jurors pre-deliberation would be returned by a jury post-deliberation and to predict a hung jury if no such verdict exists.

Additionally, mock jurors may behave differently than real world judges. The sample of mock jurors with the education and income levels of judges in this study is small, making it hard to draw clear conclusions about whether bias among them differed significantly from bias among mock jurors as a whole. Moreover, there are ways in which real world judges differ even from mock jurors with similar demographic characteristics. For example, judges are repeat players in the criminal justice system. As a result, they can compare defendants and facts across cases. Nevertheless, there is some evidence that judges generally suffer from similar biases as jurors (Guthrie, Rachlinski & Wistrich 2007; Rachlinski & Johnson 2009; Rachlinski, Wistrich, & Guthrie 2015; Wistrich, Rachlinski, & Guthrie 2015).

Even acknowledging these potential limitations, this Chapter is one of the first studies to provide evidence about the effect of defendant race on nonviolent criminal case outcomes and the first to examine the possible mechanisms through which race could be influencing these case outcomes. The results clearly indicate that the effect of defendant race on case outcomes is case-specific and can vary based on mock juror demographics. In some specifications, strong effects of race appeared in all of the case scenarios. Other models revealed that effects can depend on either the crime charged, the evidence presented, or both. In the future, experimental vignette studies examining the effect of defendant race on case outcomes should use a wider range of crimes and types of evidence.

The fact that jury nullification is at least partially responsible for the effects of defendant race on verdict decisions, and may be responsible for the effect of defendant race on sentence length among mock jurors who acquitted the defendant, complicates the policy implications of the results. There is a hearty debate among legal scholars about the appropriateness of race-based jury nullification that corrects for perceived injustices at other stages of the criminal justice system (Butler 1997; Leipold 1996). Some scholars argue that it is dangerous to allow race to influence case outcomes—no matter the reason (Leipold 1996). Others argue that race-based jury nullification has an important role to play in correcting for the very real racial prejudices that exist in the criminal justice system, and nonviolent cases are the most appropriate forum (Butler 1997). While a complete discussion of this debate is beyond the scope of this paper, it is important to note that no matter where one comes out on the issue of race-based jury nullification, it is clear that the best alternative world is one in which no race-based jury nullification is necessary to correct for biases at other stages of the criminal justice system. Not only would correcting for racial biases at other stages of the criminal justice system such as

policing, charging, and plea bargaining be a good result in and of itself, but this research suggests that it could have the additional positive impact of eliminating the effect of race in non-violent property crime cases by increasing individuals' faith in the justice system.

Finally, even though the overarching results illustrate that white defendants are worse off in nonviolent property crime cases, some negative black stereotypes are creeping into conviction and sentencing decisions (albeit, they are outweighed here by other biases against white defendants). Subjects view black defendants as having higher propensities to commit crime, such that black defendants receive lower presumptions of innocence from mock jurors who think the defendant's propensity to commit crime is important. Black defendants are also punished more severely for a given perceived recidivism likelihood, often receiving significantly longer sentences than white defendants from mock jurors who convicted. Both of these effects are clearly unconstitutional—neither a defendant's sentence nor the presumption of innocence should be affected by race in this manner. Future research examining whether these effects are conscious and whether they can be corrected with things like specialized jury instructions and judicial training can help pinpoint the most effective policies to eliminate these biases.

CHAPTER TWO: THE EFFECTS OF DEFENDANT DRESS AND SHACKLING ON NONVIOLENT CRIMINAL CASE OUTCOMES

I. Introduction

It is often presumed that a defendant's physical appearance at his or her criminal trial or sentencing hearing will matter. Lawyers advise clients not only about how to dress, but also about things like whether to chew gum and how to interact with family members in the audience (Broker & Reddington 2014). Public defender's offices often provide clients who cannot afford formal attire with a suit or dress for court (Bischoff 2018). At times, defendants have even petitioned the court for protections related to their physical appearance, relying primarily on Due Process guarantees to a fair trial. These guarantees includes the right to be presumed innocent and have one's guilt established by probative evidence beyond a reasonable doubt, and at times, the Supreme Court has interpreted this right to give defendants control over their physical appearance at trial. In *Estelle v. Williams*, the Court held that a defendant cannot be forced to appear at a jury trial in prison garb,³⁷ and in *Deck v. Missouri*, the Court held that defendants should only be shackled at a jury trial upon a finding by the court that individual circumstances—usually a risk of flight or violence—warrant the restraints.³⁸

Yet the Court's decisions in this area have not provided clear rules for lower courts to apply. Recognizing the difficulty of defining the scope of the presumption of innocence, the Court has noted that "the actual impact of a particular practice on the judgment of jurors cannot always be fully determined," and "courts must do the best they can to evaluate the likely effects of a particular procedure, based on reason, principle, and common human experience."³⁹

³⁷ 425 U.S. 501 (1976).

³⁸ 544 U.S. 622 (2005).

³⁹ *Estelle v. Williams*, 425 U.S. 501, 504 (1976).

Unsurprisingly, “reason, principle, and common human experiences” have not always resulted in consistent lower court rulings. Courts have divided over how to apply existing precedents to new technologies and unique circumstances, and courts remain divided over whether the defendant’s rights are the same during proceedings in front of a judge and proceedings before a jury.

Despite common assumptions about the importance of courtroom appearance, the need for a well-founded principle to guide courts in protecting the right to a presumption of innocence, and well-established interest in the empirical effects of beauty, relatively little prior research has examined the effect of defendant appearance on criminal case outcomes. This Chapter contributes to the existing knowledge on appearance discrimination in criminal court using an experimental vignette study to examine the effects of clothing and shackling. In the context of a nonviolent property crime, the results confirm the Court’s assumptions about jury bias *Estelle*. Certain subgroups of mock jurors convict defendants in institutional garb at higher rates than defendants in civilian clothing. The results also provide new information about the effect of the defendants’ formality in terms of non-prison attire. Defendants in institutional garb receive significantly longer prison sentences than defendants in casual civilian clothing. In contrast, the results indicate that shackling usually only affects case outcomes when considered in combination with defendant dress.

The fact that the vast majority of cases are now resolved through plea bargaining⁴⁰ in no way diminishes the importance of studying appearance discrimination at trial and sentencing. The bargaining process itself is influenced by the parties’ expectations about what would happen at trial, and because defendants have historically received more protections at trial than any other

⁴⁰ In 2014, 2.6 percent of federally adjudicated defendants received a bench or jury trial (Bureau of Justice Statistics 2017). In 2009, 2 percent of state felony convictions in the 75 largest counties in the United States resulted from a bench or jury trial (Bureau of Justice Statistics 2013).

stage of the criminal process, policy changes are likely to occur there first. Additionally, judges often retain significant discretion when sentencing defendants who have plead guilty, so the sentencing results presented here have implications that reach beyond the small subset of cases that proceed to trial.

II. Legal Background

The first time that the U.S. Supreme Court spoke on an aspect of defendant appearance outside of race was *Illinois v. Allen* in 1970.⁴¹ In that case, a pro se defendant acted in an unruly and uncontrollable manner during trial, and the judge had him removed several times, allowing the trial to proceed in his absence. On appeal, the defendant challenged his conviction based on the Confrontation Clause. While the precise holding of the Court was unrelated to the issue of shackling, the Court reasoned that “binding and gagging” the defendant would have been one constitutional way that the judge could have dealt with the defendant’s unruliness.⁴² However, the Court also recognized that “[n]o person should be tried while shackled and gagged except as a last resort.”⁴³ It justified this dicta by speculating that the restraints would likely bias the jury against the defendant, limit the defendant’s ability to assist his counsel, and harm the dignity of the courtroom.

In 1976, the Court applied the logic from *Allen* in *Estelle*, a case in which the defendant was forced to appear at trial in identifiable prison garb.⁴⁴ The Court held that defendants can never be forced to appear before a jury in identifiable prison clothing because, unlike in the shackling context, there is never a compelling state reason for making them do so. Nonetheless,

⁴¹ *Illinois v. Allen*, 397 U.S. 337 (1970).

⁴² *Id.* at 344.

⁴³ *Id.*

⁴⁴ *Estelle*, 425 U.S. 501.

in 1986, the Court declined to extend the logic of *Estelle* to prevent uniformed officers from sitting behind the defendant in the front row at a trial. While “[o]ne accused of a crime is entitled to have his guilt or innocence determined solely on the basis of evidence introduced at trial ... [t]his does not mean ... that every practice tending to single out the accused from everyone else in the courtroom must be struck down.”⁴⁵ Since there are many reasons for armed guards to be present in a courtroom, and the jury would not necessarily associate the uniformed officer with the defendant being tried, the Court decided that a showing of actual prejudice was necessary to prove a Due Process violation.

Finally, in 2007, the Court heard a case that directly presented the issue of whether indiscriminate shackling of adult criminal defendants is permissible under the Due Process Clauses of the Fifth and Fourteenth Amendments. In *Deck*, the Court held, “The Constitution forbids the use of visible shackles during the penalty phase [of a capital case], as it forbids their use during the guilt phase [of any criminal trial], unless that use is justified by an essential state interest—such as the interest in the courtroom security—specific to the defendant on trial.”⁴⁶ The Court based this decision on the same three risks discussed in *Allen*: (1) shackles would influence the jury to think that the defendant is dangerous and thereby undermine the presumption of innocence, (2) shackles would interfere with the right to effective assistance of counsel, and (3) shackles would decrease the dignity of the courtroom. When a court orders shackling in front of a jury without adequate justification, the Court held that the “defendant need not demonstrate actual prejudice.... The state must prove beyond a reasonable doubt that the shackling error did not contribute to the verdict obtained.”⁴⁷

⁴⁵ *Holbrook v. Flynn*, 475 U.S. 560, 567 (1986).

⁴⁶ *Deck v. Missouri*, 544 U.S. 622, 624 (2005).

⁴⁷ *Id.* at 635.

Courts have split on the issue of whether the Due Process protections outlined in *Deck* extend to bench trials and other judicial proceedings. The Supreme Courts of Illinois (considering shackling in the context of an adjudicatory hearing and sentencing in a juvenile case) and California (considering shackling in the context of a preliminary hearing) have applied the prohibition on indiscriminate shackling to proceedings before a judge, relying primarily on the defendant’s need to communicate with counsel and the potential of the shackles to undermine the dignity of the courtroom.⁴⁸ The highest court of New York (considering shackling in the context of a bench trial) has also applied the prohibition to proceedings before a judge, relying instead on the potential for judges to be biased against the defendant in the same manner as jurors.⁴⁹ “Judges are human, and the sight of a defendant in restraints may unconsciously influence even a judicial factfinder.”⁵⁰

Recently, the Ninth Circuit overruled its previous decision declining to extend the protection against indiscriminate shackling to proceedings before a judge,⁵¹ becoming the first federal circuit court to join these state courts.⁵² In holding that indiscriminate shackling was improper at any court proceeding, whether it be in front of a judge or jury, the court relied heavily on both the defendant’s right to a presumption of innocence and the dignity of the courtroom.⁵³ Regarding the presumption of innocence, the court noted that “[i]t includes the perception of any person who may walk into a public courtroom, as well as those of the jury, the judge and court personnel. A presumptively innocent defendant has the right to be treated with

⁴⁸ *People v. Fierro*, 821 P.2d 1302 (Cal. 1991); *In Re Staley*, 67 Ill.2d 33 (1977).

⁴⁹ *People v. Best*, 979 N.E.2d 1187 (N.Y. 2012).

⁵⁰ *Id.* at 1189.

⁵¹ *United States v. Howard*, 480 F.3d 1005, 1012–14 (9th Cir. 2007) (“[A] judge in a pretrial hearing presumably will not be prejudiced by seeing the defendants in shackles.”).

⁵² *United States v. Sanchez-Gomez*, 859 F.3d 649 (9th Cir. 2017) (en banc).

⁵³ *Id.*

respect and dignity in a public courtroom, not like a bear on a chain.”⁵⁴ The dissent argued that the case should not have been decided because the issue was moot, and, on the merits, since there was no jury and no evidence that the restraint was intended to serve as a punishment, the test should simply be whether requiring detainees to wear the restraints “is reasonably related to a legitimate governmental objective.”⁵⁵ The Supreme Court granted certiorari and recently vacated the Ninth Circuit’s opinion on the basis that the case was moot, leaving the issue of whether indiscriminate shackling is valid during judicial proceedings undecided.⁵⁶

Prior to the Ninth Circuit’s decision, the Second Circuit and the Eleventh Circuit declined to extend the prohibition on indiscriminate shackling to sentencing proceedings before a judge.⁵⁷ In reaching their decision, the Second Circuit explicitly stated that judges are not likely to experience the same biases as jurors. “We presume that where, as here, the court defers without further inquiry to the recommendation of the Marshal’s Service that a defendant be restrained at sentencing, the court will not permit the presence of the restraints to affect its sentencing decision.”⁵⁸ The Eleventh Circuit avoided discussing the issue of bias, basing its opinion primarily on historical references indicating that the right to be free of restraints is only applicable at trial.⁵⁹ The Ninth Circuit heavily criticized this historical account in *Sanchez-Gomez*.⁶⁰

⁵⁴ *Id.* at 661.

⁵⁵ *Id.* at 683 (Ikuta, J., dissenting).

⁵⁶ *United States v. Sanchez-Gomez*, 138 S. Ct. 1532 (2018).

⁵⁷ *United States v. LaFond*, 783 F.3d 1216 (11th Cir. 2015); *United States v. Zuber*, 118 F.3d 101 (2d Cir. 1997).

⁵⁸ *Zuber*, 118 F.3d at 103–105.

⁵⁹ *LaFond*, 783 F.3d 1216.

⁶⁰ 138 S. Ct. 1532.

III. Theoretical Foundations

As discussed in Chapter One, a number of theorists have developed mathematical models to analyze and predict the decisionmaking process of judges and jurors. A judge or juror reaching a verdict at trial in a criminal case should begin by presuming that the defendant is innocent. In other words, before any evidence is presented, the judge or juror should believe that there is a very small probability (approaching zero) that the defendant is guilty.⁶¹ As evidence is presented, the judge or juror weighs the credibility and sufficiency of the evidence and updates his or her belief about the defendant's likelihood of guilt. A number of mechanisms for this updating process have been proposed, perhaps the most widely accepted of which is often referred to as the story model (Pennington & Hastie 1991).⁶² Under this model, jurors assign meaning to trial evidence through the development of competing stories (Pennington & Hastie 1991). These stories are then judged based on pre-existing world knowledge to eliminate all but the "best" story (Pennington & Hastie 1991).

Whatever the precise mechanism for processing the information presented at trial, the judge or juror ultimately applies the jury instructions to their final belief about the likelihood of the defendant's guilt in order to reach a verdict decision. There are two main ways to conceptualize the ultimate verdict decision task. First, factfinders may translate the beyond a reasonable doubt standard of proof into a concrete percentage of guilt, such as 95 percent. If their post-evidence assessment of guilt is above that threshold, they will convict. Otherwise, they will not. The burden of proof threshold could be influenced by perceptions of the costs of a wrongful

⁶¹ Even absent appearance bias, prior beliefs may stray from zero based on the difficulty of presuming innocence. For example, judges or jurors may improperly consider the indictment as evidence of guilt. A similar model of jury decisionmaking is discussed and tested in Ostrom et al. (1978).

⁶² Alternative models include information integration models, Bayesian models, Poisson models, sequential weighting models, and others (Pennington & Hastie 1981).

conviction and wrongful acquittal generally, but under this conceptualization, the threshold will be consistent across cases absent the influence of bias. Alternatively, the judge or juror may reach his or her verdict by directly comparing the costs of a wrongful conviction and wrongful acquittal. The weights that they assign to each cost will depend on their post-evidence assessment of the defendant's guilt, and the size of the costs will depend on their personal beliefs and interpretations of the legal instructions. Under this conceptualization, the burden of proof threshold may vary between cases, even absent bias. This is the conceptualization that best aligns with the story model of jury decisionmaking. Under that model, the decisionmaker must decide whether the "best" story contains all the attributes necessary for a guilty verdict beyond a reasonable doubt (Pennington & Hastie 1991).⁶³ This decision will depend on the decisionmakers prior ideas about crime categories, intuition, and interpretation of the legal instructions (Cheng 2013; Pennington & Hastie 1991).

Similar to race, discussed in Chapter One, a judge or juror who is not biased by the defendant's appearance will not allow his or her pre-evidence belief, belief updating, or verdict determination to be influenced by the defendant's appearance. In contrast, a biased judge or juror may—consciously or unconsciously—be influenced by the defendant's appearance in a number of ways.

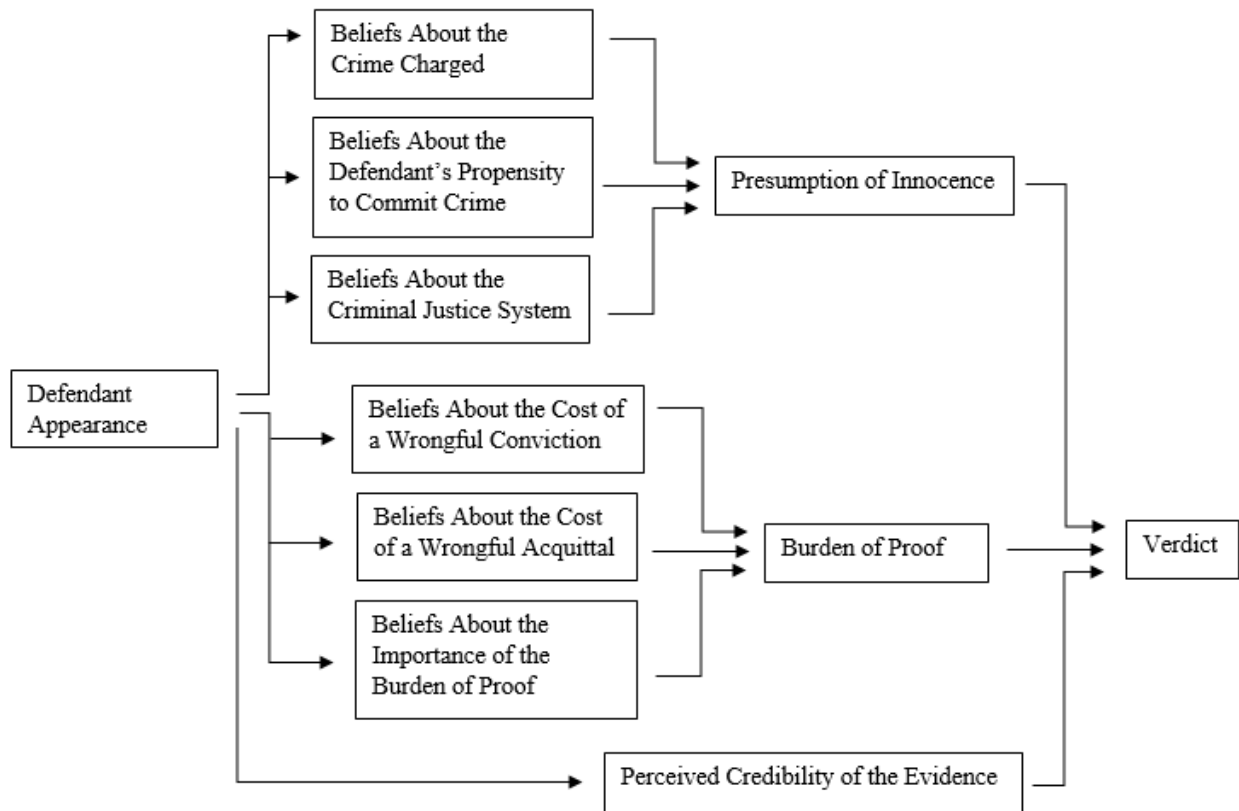
Before any evidence is presented, a biased judge or juror may allow stereotypes to influence his or her belief about the probability of the defendant's guilt. Whatever the source of the stereotype, it will result in different initial probabilities of guilt for different defendants, in contravention of the presumption of innocence instruction. A biased judge or juror may also be

⁶³ Alternatively, jurors might convict if they believe that the prosecutor's story is the most convincing and acquit otherwise, failing to consider the burden of proof in a third step. If this is the case, the burden of proof results presented here may not correlate with actual verdict decisions.

influenced—consciously or unconsciously—by the defendant’s appearance during the presentation of evidence. He or she may find certain types of evidence more credible based on the defendant’s appearance or weigh the evidence differently. Finally, a biased judge or juror may—consciously or unconsciously—allow the defendant’s appearance to affect his or her interpretation of the burden of proof. Beyond a reasonable doubt may be a more demanding standard for defendants with a certain appearance, or the costs of a wrongful conviction and wrongful acquittal may differ based on the defendant’s appearance.

As Figure 1 illustrates, all of these forms of appearance bias combine to influence aggregate differences in verdicts, and they may work in opposite directions. For example, it is theoretically possible that judges and jurors associate shackling with a propensity to commit crime and exhibit bias against shackled defendants even before viewing any evidence. However, judges and jurors may also find the evidence presented against shackled defendants less credible because the defendant was incarcerated and had less opportunity to find and present countervailing evidence. In this case, if their application of the burden of proof is not influenced by shackling, it is theoretically unclear whether shackled defendants will be found guilty more or less often than unshackled defendants.

Figure 1. Mechanisms Through Which Appearance May Impact Verdict Decisions



Judges and juries exercise much more discretion when assigning sentences than when determining guilt or innocence.⁶⁴ When a sentence is assigned, the defendant has already been found guilty beyond a reasonable doubt, and the goals are to rehabilitate the defendant, deter future crime, incapacitate the defendant while he or she is a danger to the community, and obtain retribution. The judge or jury will weigh these various goals based on their personal preferences and beliefs about the defendant and determine the optimal sentence within the statutorily permissible range. In doing so, the judge or jury may consider a wide range of factors, including the defendant's probability of reoffending. Oftentimes, judges are provided with statutory

⁶⁴ Most sentencing decisions are made by judges, but some sentencing decisions are left to juries, even outside the capital murder context. States that allow for jury sentencing in noncapital cases are Arkansas (ARK. CODE ANN. § 16-97-101), Kentucky (KY. REV. STAT. ANN. § 532.055), Missouri (MO. ANN. STAT. § 557.036), Oklahoma (OKLA. STAT. ANN. tit. 22 § 926.1), Texas (TEX. CODE CRIM. PROC. ANN. art. 37.07, § 2), and Virginia (VA. CODE ANN. § 19.2-295).

sentencing guidelines to help them weigh various factors in coming to their decision. Such guidelines are typically advisory and can be very complex.

The wide discretion given to decisionmakers at the sentencing stage makes it legally permissible for judges and jurors to consider defendant dress and shackling when assigning a sentence.⁶⁵ For example, a judge or juror may see that a defendant is shackled, assume that this is based on a determination by the sheriff's deputies that the defendant is dangerous, and issue a higher sentence in response. A judge or juror may see that a defendant is in casual rather than formal attire, assume that this is based on a lack of respect for the court, and issue a higher sentence in response. Neither of these sentencing decisions are "biased" in the sense that they are unconstitutional. Nevertheless, policymakers may prefer sentences to be based on factors other than defendant appearance. For that reason, understanding the effects that defendant dress and shackling have on sentencing decisions is important in order to inform policy decisions relating to the defendant's ability to control his or her appearance.

IV. Existing Literature

There is almost no existing literature on the effects of defendant dress and shackling, but appearance literature generally—especially literature on the effect of beauty—is much more common, both in and out of the criminal context. In their seminal paper on beauty, Hamermesh and Biddle (1994) find that plain looking people earn less than the good-looking. This effect holds for both men and women, and is largely independent of occupation, indicating that it may be the result of taste-based employer discrimination—i.e., discrimination stemming from

⁶⁵ Capital sentencing proceedings are an exception to this rule. In *Deck*, the Court held that indiscriminate shackling was not permissible at the penalty phase of a capital murder trial even though the defendant is no longer presumed innocent of the crime, because the need for an accurate and unbiased determination of the appropriateness of the death penalty is so critical. 544 U.S. at 632.

decisions made by persons deriving disutility from interacting with less attractive individuals (Becker 1957). In a related paper, Mobius and Rosenblat (2005) decompose the source of the beauty premium in a labor market experiment. They find that attractive workers are more confident, and confidence increases wages; for a given level of confidence, attractive workers are (wrongly) considered to be more able by employers; and controlling for confidence, attractive workers have communication and social skills that raise their wages when they interact with employers. Other studies confirm the wage and employment effect of beauty both in the United States (Fletcher 2009) and other countries (Galarza & Yamada 2016), as well as specifically for private-sector attorneys (Biddle & Hamermesh 1998) and managers (Frieze, Olson, and Russell 1991).

Beauty has also been shown to impact happiness (Gupta, Etcoff, & Jaeger 2015; Hamermesh & Abrevaya 2013), political elections (Berggren, Jordahl, & Poutvaara 2017; Jones & Price 2017; Leigh & Susilo 2009), charitable giving (Cryder, Botti, & Simonyan 2017; Price 2008), instructor ratings (Hamermesh & Parker 2005), profits (Rule & Ambady 2008), academic performance (Hernandez-Julian & Peters 2017), and propensity to commit crime (Mocan & Tekin 2010). Hamermesh, Meng, and Zhang (2002) find that women's spending on beauty-enhancing goods and services increases their perceived beauty, which in turn impacts their wages. Nevertheless, such purchases pay back in earnings no more than 15 percent of total expenditures.

In the criminal context, most experimental studies find that defendant beauty impacts guilt (Abwender & Hough 2001; Desantis & Kayson 1997; Efran 1974; Leventhal & Kreate 1977; MacCoun 1990; Patry 2008) and sentencing (Abwender & Hough 2001; Efran 1974; Friend & Vinson 1974; Leventhal & Kreate 1977; Piehl 1977; Sigall & Ostrove 1975; Smith &

Hed 1997) decisions. Typically, more attractive defendants are better off, but a meta-analysis by Mazzella and Feingold (1994) finds that the effect of beauty on both guilt determinations and sentencing is small, and the sentencing effects of beauty differ based on the crime charged. Attractive defendants receive less punishment for robbery, rape, and cheating; more punishment for negligent homicide; and the same amount of punishment as unattractive defendants for swindling. Researchers have also studied other aspects of appearance and found that eyeglasses reduce the likelihood of a guilty verdict (Brown, Henriquez, & Groscup 2008), and facial tattoos increase the likelihood of a guilty verdict but not the severity of sentencing (Funk & Todorov 2013). Obesity can also lead to increased perceptions of guilt for female defendants (Schvey et al. 2013).

There are two main ways that attractiveness is measured in the existing literature. The first is by presenting a group of individuals (in the experimental context this group is often different from the experiment participants) with a picture and asking them to rate attractiveness on a Likert scale (Biddle & Hamermesh 1998; Efran 1974; Leventhal & Kreate 1977; MacCoun 1990; Mobius & Rosenblat 2005; Price 2008; Rule & Ambady 2008). The second, used only in the experimental context, involves signaling beauty through a written case description (Friend & Vinson 1974). For example, the word “attractive” or “unattractive” might be used to describe the defendant in a written description of the case.

Attire is one aspect of appearance that has been shown to correlate with attractiveness (Jacobsen & Berger 1974). It is particularly important to study in the context of the criminal justice system, not only because the Supreme Court has singled it out for examination,⁶⁶ but also because it is an aspect of appearance that is often within the control of the defendant. There are

⁶⁶ Estelle v. Williams, 425 U.S. 501, 504 (1976).

two existing studies that directly examine the impact of defendant dress on criminal justice outcomes. Fontaine and Kiger (1978) conduct an experimental vignette study in which respondents watch a videotape of a first-degree murder trial. When it is the defendant's turn to testify, the video displays a still frame of the defendant in one of four conditions: (1) wearing personal clothing (sport coat, tie, and slacks) and not accompanied by an armed guard, (2) wearing personal clothing (sport coat, tie, and slacks) and accompanied by an armed guard, (3) wearing institutional clothing and not accompanied by an armed guard, or (4) wearing institutional clothing and accompanied by an armed guard. Defendants in conditions two and three are significantly more likely to be found guilty and are given significantly higher sentences than defendants in conditions one and four, indicating some bias against defendants who are either in institutional dress *or* accompanied by an armed guard but no bias against defendants who are both in institutional dress *and* accompanied by an armed guard. The authors explained this result by noting participants report feeling most sympathetic towards defendants who are in institutional dress and accompanied by an armed guard.

Jacobsen and Berger (1974) conduct an experimental vignette study in which respondents are asked to assign a sentence between one and twenty-five years to a defendant who has been convicted of negligent vehicular homicide. Photographs of the same male model are attached to each questionnaire. In the "high attractiveness" condition the defendant is dressed in a coat and tie, wearing glasses, with his hair neatly combed. In the "low attractiveness" condition he is wearing a t-shirt and jeans with his hair combed back in a loose style. The authors also manipulate the defendant's repentance level. They do not find a significant effect of defendant appearance on sentence length.

Although not the principal subject of their study, Conley, Turnier, and Rose (2000) also examine the effect of defendant dress indirectly using an experimental vignette study that varies the socioeconomic status and race of the defendant in a videotape of a vandalism trial. Upper-class defendants wear blue blazers, blue shirts, and striped ties. They testify that they are engineering students and live with both parents; their mothers work as physicians and their fathers own a business. Lower-class defendants wear khaki pants and open-collared shirts. They testify that they work as apprentice mechanics in a tire and muffler shop and live in a trailer park with single mothers who work as nurses' aides in a nursing home. The authors find that socioeconomic status does not have a significant effect on determinations of guilt. No previous studies have looked at the effects of shackling.

Additionally, no previous literature has examined how the effect of defendant dress and shackling on verdict determinations and sentencing differs based on the demographic characteristics of the decisionmaker, but there are some studies that examine how the effect of defendant beauty interacts with mock juror characteristics. Abwender and Hough (2001) find that female mock jurors treat attractive female defendants more harshly than unattractive female defendants, and male mock jurors treat attractive female defendants less harshly than unattractive female defendants. Similarly, Dumas and Teste (2006) find that female mock jurors sentence attractive male defendants to more jail than unattractive male defendants, and male mock jurors do the opposite. In contrast, Sigall and Ostrove (1975) find no differences in the effect of defendant attractiveness by mock juror gender, and Nemeth and Sosis (1973) find that a sample of working class, conservative mock jurors assigned more severe penalties to convicted defendants who are white and unattractive as compared to a wealthier and more liberal sample of mock jurors.

There is also literature finding effects of juror demographics on case outcomes more generally, but the results are sometimes mixed. Mills and Bohannon (1980) find that women convict more often than men, with black female mock jurors convicting at the highest rate; jurors between the ages of 18 and 25 and jurors between the ages of 50 and 65 convict at the highest rates; and the frequency of guilty verdicts decreases with education. Conversely, Baldwin and McConville (1980) find no effect of juror age, gender, or social class on verdicts. Culhane, Hosch, and Weaver (2004) find that mock jurors who have previously been victims of property crimes are more likely to convict a property crime defendant, while mock jurors who have previously been victims of violent crimes are not, and Miller et al. (2011) find that mock jurors are less likely to convict defendants with their same religion.

Finally, there are some juror characteristics that, despite not having been empirically examined, are often assumed to correspond to juror decisionmaking. Rural jurors are often assumed to be more “backwards thinking” than urban jurors (Bassett 2006), and jurors with a criminal history are often struck from juries for cause or excluded entirely from jury pools (Kalt 2003; Salyers 1999). Nevertheless, a study by Binnall (2017) indicates that convicted felons are particularly committed to serving as conscientious jurors. Overall, this literature suggests that the effects of defendant dress and shackling could differ based on a number of juror demographics.

V. Study Design

This Chapter uses data from the same pair of experimental vignette studies as Chapter One to examine the impact of defendant dress and the presence or absence of shackles on verdict and sentencing decisions in four nonviolent criminal case scenarios. Experimental vignette studies are particularly valuable in this context because the Supreme Court has banned the indiscriminate use of shackling at jury trials and prohibited the state from forcing defendants to

appear at jury trials in institutional clothing.⁶⁷ Observational data on the outcomes of cases where the defendant is in institutional clothing are sparse and endogenously attributable to defendant choice to remain in institutional dress. Similarly, observational data on the outcomes of cases where the defendant was shackled at trial are likely to suffer from extreme selection bias because the sample will contain a gross overrepresentation of cases where the judge made a particularized finding about the defendant's dangerousness or propensity to flee. Even testing the effect of formal versus casual clothing choices using real world data would be problematic due to selection bias and the number of confounding variables—including socioeconomic status—that would have to be controlled for to avoid omitted variables bias.

The two surveys were designed using Qualtrics technology and run using the mTurk platform. In total, 8,070 respondents participated in the first survey, which included two separate vignette studies⁶⁸ and took about fifteen minutes to complete. Participants were only allowed to complete the survey once⁶⁹ and were compensated with \$1.50 upon successful completion. All subjects were over eighteen, resided within the United States, and spoke English. In general, the mTurk sample is younger, more educated, and more likely to be employed than the U.S. population. The mTurk sample also has a slight overrepresentation of women, an underrepresentation of racial and ethnic minorities, and is more urban than the United States population.⁷⁰

Additionally, 2,561 individuals participated in a second survey, and they were compensated \$0.25 for about 2.5 minutes of their time. Again, all of the subjects were over

⁶⁷ *Deck v. Missouri*, 544 U.S. 622 (2005) (banning indiscriminate shackling of adult criminal defendants); *Estelle v. Williams*, 425 U.S. 501 (1976) (providing adult defendants the right not to be seen by the jury in prison garb).

⁶⁸ The other vignette study, which was not related to this project, involved employee hiring scenarios.

⁶⁹ This requirement is easily enforced since each participant must be signed up as an mTurk worker and each worker is assigned a unique worker ID.

⁷⁰ All of these demographic characteristics are typical of mTurk samples (see Ginther et al. 2014; Hersch & Shinall 2017).

eighteen, resided within the United States, and spoke English. Demographic characteristics of this sample are provided in Table 1, alongside those of the first mTurk sample and the U.S. population according to the 2016 Census. For the most part, the demographic characteristics of the second sample mirror those of the first sample. One notable difference is a more representative proportion of Hispanic individuals (bringing the percent closer to that of the U.S. population). Additionally, as compared to the first study, a greater percentage of the second study respondents live in urban areas, identify as Republican, identify as religious, have prior jury experience, and trust the justice system. While these differences are large, they are not a major concern. The samples are drawn from the same population of mTurk workers, and they are rarely combined for analysis in this Chapter.

Other researchers have found that mTurk workers are comparable to other validated panels (Kuziemko et al. 2015), and have used mTurk samples for research on jury decisionmaking in the criminal context (Ingriselli 2015; Ginther et al. 2014), reducing concern that mTurk workers give systematically different answers than nationally representative samples.

Table 1. Subject Demographics

Characteristic	Second mTurk Sample	First mTurk Sample	U.S. Population (2016)
Female	54.5%	55.3%	51.1%
Age	35.3	35.4	41.8
US Native	94.4%	95.0%	86.8%
Hispanic	12.1%	8.1%	14.4%
White	79.2%	79.0%	76.4%
Black	8.5%	8.7%	10.2%
Asian	6.7%	6.3%	5.4%
Married	46.6%	41.1%	42.6%
B.A. or Higher (If 25+)	60.7%	57.7%	31.4%
Employed	84.9%	82.3%	46.2%
Median Household Income (\$2017)	\$52,500	\$52,500	\$58,844
Urban	39.5%	33.7%	21.8%
Northeast	17.0%	19.1%	17.7%
Midwest	19.3%	21.3%	21.5%
South	42.6%	38.5%	37.3%
West	20.9%	20.9%	23.4%

Democrat	40.2%	43.3%	N/A
Republican	30.5%	22.4%	N/A
Religious	34.7%	23.6%	N/A
Prior Jury Service	16.7%	13.6%	N/A
Criminal History	18.8%	17.0%	N/A
Criminal Victimization	41.9%	41.4%	N/A
Distrusts Justice System	31.9%	44.0%	N/A

Notes: Mean values for the U.S. population are calculated from the 2016 U.S. Census Bureau ACS sample. Age is calculated for employed persons only; race values are calculated for individuals who report only a single race. Individuals who attend church at least once a month are considered religious. Individuals who have a prior criminal conviction or have been arrested are considered to have a criminal history. Individuals are considered to distrust the justice system if they strongly disagreed or somewhat disagreed with the statement that the criminal justice system works well in the United States.

In the first survey, as discussed in more detail in Chapter One, each respondent was randomly assigned to see either an eyewitness or motive variation of the vandalism scenario, and then they saw the opposite variation of the arson scenario. At the beginning of each scenario, subjects were prompted that they had been selected to serve as a juror in a criminal trial. Then they read a written description of the trial, including witness testimony and attorney arguments.

Most importantly for this experiment, each scenario was accompanied by a picture of a defendant. Pictures of two different white defendants and two different black defendants were used. The clothing of the defendant was randomly selected to be either a navy suit and tie (formal), jeans and a white t-shirt (casual), or an orange jumpsuit (institutional). In addition, the defendant was randomly depicted with or without shackles. The shackles consisted of handcuffs and anklecuffs connected by a chain. However, only the handcuffs and the chain going towards the floor were visible because the pictures were cut off above the ankles to prevent shoe style differences from impacting the results. All of the pictures of the defendants were taken in front of the same white wall, and participants were instructed that the picture depicted how the defendant appeared in court on the day of trial. No participants saw the same defendant twice. Figure 2 contains three example defendant pictures—one where the defendant is wearing formal

attire without shackles, one where the defendant is wearing casual attire without shackles, and one where the defendant is wearing institutional attire with shackles.

Figure 2. Example Defendant Photographs



Each scenario concluded with jury instructions about the burden of proof, the standard of proof, the presumption of innocence, the elements of the crime charged, and the right to remain silent. Respondents were asked to find the defendant either guilty or not guilty, rate their confidence in that guilt determination on a five point Likert scale, rate the believability of each witness and each party's version of events on a five point Likert scale, recommend a sentence, rate their confidence in the sentence determination on a five point Likert scale, and rate the likelihood that the defendant would reoffend in the next five years if acquitted on a five point

Likert scale. The questions were asked in this order on separate screens, and respondents could not go back and change a previous answer.

For the sentencing determination, participants were asked to assume the role of a judge and, regardless of their answers to the previous questions, assume the defendant had been found guilty. They were told that the defendant's criminal history consisted of one misdemeanor theft conviction as a juvenile, and their answers were restricted to between one and five years (inclusive) for vandalism and between one and ten years (inclusive) for arson. To simplify the decision, there was not an option to impose a non-incarceration sentence, such as probation. At the end of the survey, respondents were asked a variety of demographic questions. A complete survey outline is attached as Appendix A.

The second study, also discussed in more detail in Chapter One, was intended to gauge the effect of pre-existing appearance bias—as compared to appearance bias that develops based on the presentation of evidence—on the presumption of innocence and the burden of proof (see Figure 1). Participants were also asked to rate the attractiveness of the various defendants. More specifically, each respondent was presented with one criminal case scenario (either arson or vandalism) and then one attractiveness question about a different defendant.

The criminal case scenario began with a prompt informing the participant that they had been selected to serve on a jury in a criminal trial, informing them of the charge, and providing them with the jury instructions that they would be asked to apply after hearing the evidence. The jury instructions were the same as those used in the first survey. The prompt was accompanied by a randomly assigned picture of a defendant. The same pictures were used as in the first survey, and participants were instructed that the defendant was pictured as he appeared in court on the day of trial.

Participants were asked how likely they thought it was that the defendant was guilty, on a scale of 0 to 100 percent, before seeing any evidence. They were also asked how important (1) the crime charged, (2) the defendant's propensity to commit crime, and (3) their opinions about the criminal justice system were in assessing pre-trial guilt on a five point Likert scale. Then, participants were asked how certain they would need to be of the defendant's guilt, on a scale of 0 to 100 percent, after hearing the evidence in order to return a guilty verdict. Participants were also asked to rate how important (1) fear of a wrongful conviction, (2) fear of a wrongful acquittal, and (3) whether the prosecution had met the burden of proof would be in deciding whether to convict on a five point Likert scale.

After the criminal case scenario, participants were asked to rate the attractiveness of a different defendant on a five point Likert scale. Like the criminal case scenarios, this question included a randomly assigned picture of a defendant. The study concluded with a series of demographic questions. A complete survey outline is attached as Appendix B.

VI. Conviction Results

A. Appearance Does Not Affect Conviction Rates When Mock Jurors Are Pooled

With all of the mock jurors pooled together, there are no significant effects of defendant dress or shackling on verdict decisions. Table 2 summarizes the conviction rates for each of the case scenarios in the first study and illustrates how they differ based on the appearance of the defendant. Arson cases result in higher conviction rates than vandalism cases and eyewitness variations result in higher conviction rates than motive variations. Eighty percent of participants are either somewhat confident or very confident that they have reached the legally correct verdict. Among those who convict the defendant, about 84 percent are somewhat or very

confident about their decision, significantly more than the approximately 77 percent of participants who feel the same after acquitting the defendant.

T-tests were conducted to determine if conviction rates differ significantly based on the presence or absence of shackles alone, Pearson’s chi-squared tests were conducted to determine if conviction rates differ significantly based on defendant attire alone, and Pearson’s chi-squared tests were also conducted to determine if conviction rates differ significantly between any of the six appearance conditions, allowing for shackling and attire to interact.⁷¹ There are no statistically significant differences, indicating that mock jurors as a whole are not influenced by the defendant’s attire or the presence or absence of shackles. Figure 3 illustrates the lack of disparity in conviction rates by defendant appearance.

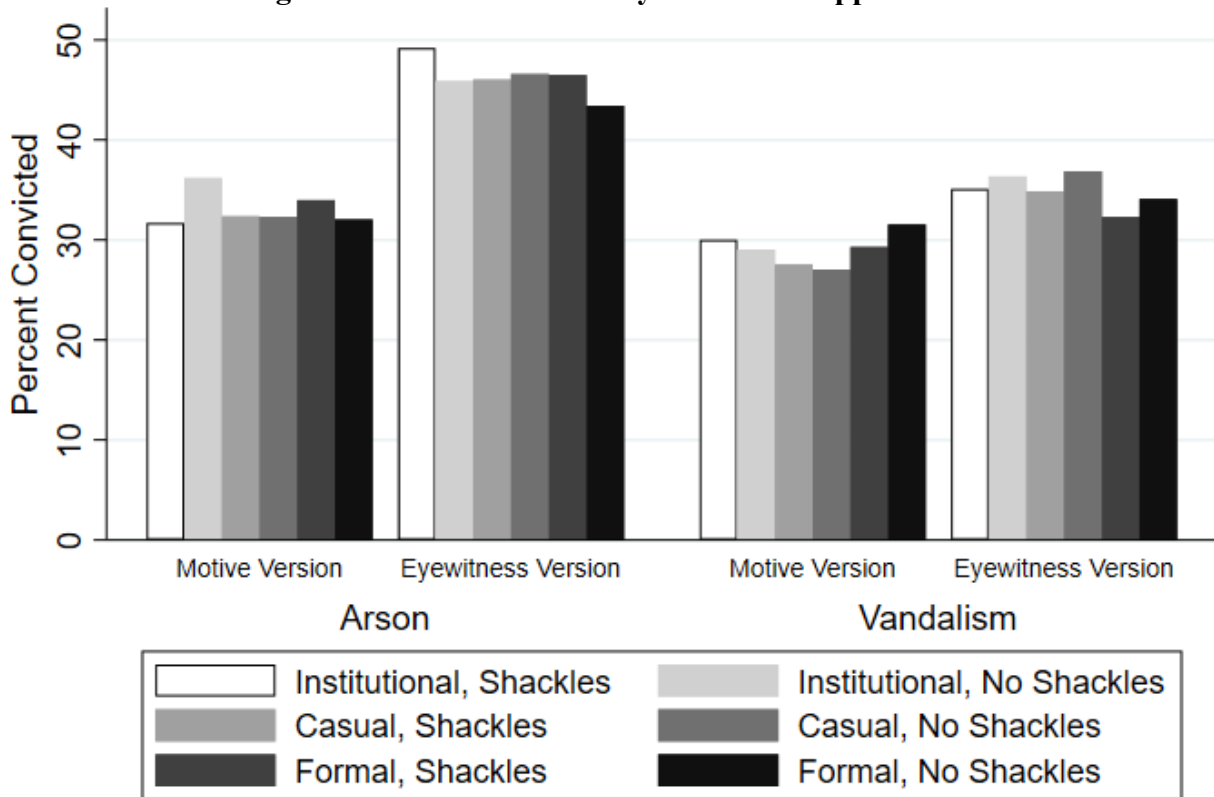
Table 2. Conviction Rates by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	35.82%	34.94%	29.08%	46.30%	33.06%
Institutional, Shackles	36.61%	35.11%	30.02%	49.20%	31.71%
Institutional, No Shackles	36.85%	36.34%	29.01%	45.88%	36.18%
Casual, Shackles	35.06%	34.88%	27.53%	46.04%	32.36%
Casual, No Shackles	35.72%	36.88%	27.01%	46.58%	32.32%
Formal, Shackles	35.62%	32.29%	29.29%	46.50%	33.98%
Formal, No Shackles	35.11%	34.10%	31.54%	43.43%	32.04%

Notes: T-tests indicate that conviction rates do not differ significantly based on the presence or absence of shackles alone. Pearson’s chi-squared tests indicate that conviction rates do not differ significantly based on the defendant’s attire alone and that conviction rates do not differ significantly by the defendant’s appearance condition (a combination of attire and shackling).

⁷¹ In essence, Pearson’s chi-squared tests determine whether the proportion of respondents returning a verdict of guilty are equivalent across the three attire conditions and the six appearance conditions. They are appropriate because both variables being examined (in this case verdict and appearance/attire) are categorical, and the observations in each appearance group are independent.

Figure 3. Conviction Rates by Defendant Appearance



The results from Chapter One illustrate that mock jurors’ verdict determinations are influenced by the race of the defendant, but appearance does not affect the conviction rates of either white or black defendants when all mock jurors are pooled. Tables 3 and 4 contain conviction rates by case scenario and defendant appearance for the subsamples of white and black defendants, respectively. Again, T-tests were conducted to determine if conviction rates differ significantly based on the presence or absence of shackles alone, Pearson’s chi-squared tests were conducted to determine if conviction rates differ significantly based on defendant attire alone, and Pearson’s chi-squared tests were conducted to determine if conviction rates differ significantly between the six appearance conditions, allowing for shackling and attire to interact. There are no significant results.

Table 3. Conviction Rates of White Defendants by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	39.14%	36.73%	31.54%	51.33%	37.12%
Institutional, Shackles	40.23%	36.11%	34.20%	55.13%	34.92%
Institutional, No Shackles	39.56%	35.93%	30.27%	52.32%	40.25%
Casual, Shackles	38.43%	37.64%	30.21%	48.90%	37.68%
Casual, No Shackles	39.42%	38.73%	30.65%	50.58%	37.28%
Formal, Shackles	39.47%	35.23%	33.22%	52.24%	36.42%
Formal, No Shackles	37.76%	36.51%	31.14%	48.31%	36.47%

Notes: T-tests indicate that conviction rates do not differ significantly based on the presence or absence of shackles alone. Pearson's chi-squared tests indicate that conviction rates do not differ significantly based on the defendant's attire alone and that conviction rates do not differ significantly by the defendant's appearance condition (a combination of attire and shackling).

Table 4. Conviction Rates of Black Defendants by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	32.54%	33.05%	26.76%	41.46%	29.02%
Institutional, Shackles	33.06%	34.14%	26.19%	43.35%	28.32%
Institutional, No Shackles	34.20%	36.77%	27.79%	40.06%	32.21%
Casual, Shackles	31.58%	32.07%	24.69%	43.13%	26.92%
Casual, No Shackles	31.78%	34.87%	23.51%	42.17%	26.84%
Formal, Shackles	32.04%	29.06%	26.11%	41.03%	31.86%
Formal, No Shackles	32.57%	31.40%	31.19%	39.32%	27.88%

Notes: T-tests indicate that conviction rates do not differ significantly based on the presence or absence of shackles alone. Pearson's chi-squared tests indicate that conviction rates do not differ significantly based on the defendant's attire alone and that conviction rates do not differ significantly by the defendant's appearance condition (a combination of attire and shackling).

B. Institutional Clothing Can Increase the Odds of Conviction and Casual Clothing Can Decrease the Odds of Conviction

Even though there is no effect of defendant dress or shackling on conviction decisions sample-wide, it is possible that certain subgroups of mock jurors are significantly influenced by defendant appearance. This could happen if the effects of defendant appearance on certain subgroups of mock jurors are cancelled out by opposite effects on other subgroups, diluted due to the small size of the subgroup, or both. In Chapter One, previous literature suggested that defendant race may have a differential effect on case outcomes based on the race and southern

residency of the mock juror. In fact, the effect of defendant race on verdict determinations does differ significantly based on the race of the mock juror.

As discussed above, no previous literature has directly examined how the effect of defendant dress and shackling on verdict determinations and sentencing differs based on the demographic characteristics of the decisionmaker. Nevertheless, studies examining how the effect of defendant beauty interacts with mock juror characteristics and studies examining the effect of juror demographics on case outcomes more generally illustrate the potential importance of heterogeneity by juror demographics. In order to explore the interaction between mock juror characteristics and the effect of defendant appearance on conviction rates, Equation 1 is estimated using a linear probability model for subgroups of mock jurors from the first study, defined by the demographic characteristics listed below.

$$Guilty_i = \alpha + \beta_1 Casual_i + \beta_2 Institutional_i + \beta_3 Shackles_i + \beta_4 Arson_i + \beta_5 Eyewitness_i + \epsilon_i \quad (1)$$

$Guilty_i$ is a dummy variable equal to one if the defendant is convicted. $Casual_i$ and $Institutional_i$ are dummy variables equal to one if the defendant is in casual and institutional attire, respectively. Formal attire is the omitted clothing category. $Shackles_i$ is a dummy variable equal to one if the defendant is shackled. $Arson_i$ is a dummy variable equal to one when the crime charged is arson, and $Eyewitness_i$ is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering on a motive. In order to maintain adequate sample sizes, all four case scenarios are pooled for this analysis. Standard errors are clustered at the individual level.

The subgroups of mock jurors are: (1) men and women; (2) black-only, white-only, and Asian-only; (3) Hispanic and non-Hispanic; (4) those ages 18-24, those ages 25-34, those ages 35-54, and those ages 55 and up; (5) those with and without a bachelor's degree; (6) those with

an income less than the median of \$52,500 per year and those with at least the median income of \$52,500 per year; (7) urban and nonurban residents; (8) Democrats and Republicans; (9) those who attend church at least once and month and those who do not; (10) those with some criminal history and those with no criminal history;⁷² and (11) those who have previously been the victim of a crime and those who have not.

Male mock jurors, Republican mock jurors, religious mock jurors, and mock jurors with no criminal history are all significantly influenced by defendant dress: They are more likely to convict a defendant in institutional attire as compared to a defendant in formal attire. The effect is strongest among Republican mock jurors who are just over 5 percentage points more likely to convict defendants in institutional attire than defendants in formal attire. Mock jurors who identify as Democrat are not significantly influenced by defendant attire but are about 2 percentage points less likely to convict a defendant that is shackled as compared to a defendant that is not shackled. These results are presented in Table 5.

Table 5. Regression Analyses of Defendant Dress and Shackling on Conviction Decisions Made by Democrat, Republican, Male, and Religious Mock Jurors and Mock Jurors Who Have No Criminal History

	Democrat	Republican	Male	Religious	No Criminal History
Casual	1.36%	2.94%	-0.01%	2.78%	0.11%
Institutional	0.99%	5.25%***	3.47%**	4.91%**	2.35%**
Shackles	-2.27%**	0.56%	0.02%	0.82%	0.08%
Arson	7.60%***	9.26%***	7.47%***	7.48%***	7.58%***
Eyewitness	8.23%***	12.04%***	8.99%***	9.18%***	9.48%***
Observations	6,716	3,482	7,216	3,714	13,198
R-Squared	0.015	0.025	0.016	0.017	0.016

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

Notes: Standard ordinary least squares regressions are used. The outcome variable is a dummy variable equal to one if the defendant is found guilty. Casual and Institutional are dummy variables equal to one if the defendant is in that attire. Shackles is a dummy variable equal to one if the defendant is shackled. Arson is a dummy variable equal to one if the crime charged is arson, and eyewitness is a dummy variable equal to one if the crime centers on the testimony of an eyewitness rather than a motive. Standard errors are clustered at the individual level.

⁷² Individuals who report having a prior misdemeanor or felony conviction or having been arrested are considered to have a criminal history.

The subgroups of mock jurors that are influenced by defendant dress overlap. Political affiliation is significantly correlated with religiosity, and men, Republicans, and those who are religious are all less likely to have a criminal history than their counterparts. However, that overlap does not mean that the effect of defendant dress is duplicative. In fact, when the sample of mock jurors who are male, Republican, religious, and have no criminal history are isolated (n=518), the effect of defendant dress is about equal to the sum of all the subgroup effects: defendants in formal attire are about 16 percentage points less likely to be found guilty than defendants in institutional attire.

While the rest of the subgroups of mock jurors are not significantly affected by defendant dress or shackles alone when making their verdict determinations, it is still possible that they are significantly impacted by certain combinations of defendant dress and shackling. In order to explore these interaction effects, Equation 2 is estimated using linear probability models for the same subgroups of mock jurors listed above:

$$Guilty_i = \alpha + \beta_1 AppearanceCondition_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \epsilon_i \quad (2)$$

Guilty is a dummy variable equal to one if the defendant is convicted, *AppearanceCondition* is vector of dummy variables equal to one for five of the six appearance conditions; formal attire with no shackles is the omitted category. *Arson* is a dummy variable equal to one when the crime charged is arson, and *Eyewitness* is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering on a motive. In order to maintain adequate sample sizes, all four case scenarios are pooled for this analysis. Standard errors are clustered at the individual level.

In addition to male mock jurors, Republican mock jurors, religious mock jurors, and mock jurors with no criminal history, three additional subgroups are affected by the interaction

of defendant dress and shackling when making conviction decisions: black mock jurors, mock jurors between the ages of 25 and 34, and mock jurors with some criminal history. Black mock jurors are only affected by defendant dress when the defendant is not shackled: They are almost 10 percentage points more likely to convict defendants in institutional attire without shackles than defendants in formal attire without shackles. Mock jurors between the ages of 25 and 34 convict defendants who are in institutional attire without shackles about 5 percentage points more often than defendants in casual attire with shackles, indicating a bias against defendants in institutional attire in combination with a preference for defendants in casual attire. Finally, mock jurors with a criminal history convict defendants in formal attire without shackles at the *highest* rate. They convict such defendants about 7 percentage points more often than defendants in institutional attire with shackles and defendants in casual attire without shackles.⁷³

Two subgroups of mock jurors that are particularly important to examine are those with characteristics similar to actual judges and jurors. Table 6 contains conviction rates by defendant appearance for mock jurors who have and have not served on a jury, mock jurors with and without the education and income levels of a typical judge, and mock jurors with and without the characteristics most common among seated jurors. As in Chapter One, based on data from the Bureau of Labor Statistics' Occupational Employment Statistics survey, individuals are considered to have the education level and approximate income level of judges if they have a doctoral or professional degree and an income of \$75,000 per year or more (Bureau of Labor Statistics 2018). Based on data from a survey administered to seated jurors in four courts by

⁷³ There are only 352 mock jurors in the first study that report having a prior felony conviction, but these individuals represent a particularly interesting subgroup of those with some criminal history as they are automatically excluded from jury service in many jurisdictions (Kalt 2003). The subgroup of mock jurors who identify as prior felons respond to defendant appearance similarly to the larger group of individuals with some criminal history: They convict defendants in formal attire without shackles at the highest rate, significantly more than defendants in institutional attire (with or without shackles) and significantly more than defendants in casual attire without shackles.

Hannaford-Agor et al. (2003) and the demographic characteristics of participants who report prior jury service in the first study, individuals are considered to have the most common characteristics of seated jurors if they have at least some college education and are employed.

Table 6. Conviction Rates by Defendant Appearance, Judicial and Juror Demographics, and Prior Jury Service

	Prior Jury Service	No Prior Jury Service	Characteristics Similar to Seated Jurors	Characteristics Distinct from Seated Jurors	Education and Income Similar to Judges	Education and Income Distinct from Judges
All Defendants	34.45%	36.04%	35.32%	37.42%	32.80%	36.06%
Institutional, Shackles	38.33%	36.36%	35.76%	39.42%	31.34%	37.03%
Institutional, No Shackles	33.16%	37.47%	36.47%	38.04%	36.41%	36.85%
Casual, Shackles	33.24%	35.34%	35.17%	34.71%	37.56%	34.74%
Casual, No Shackles	32.72%	36.13%	36.09%	34.50%	29.32%	36.19%
Formal, Shackles	35.13%	35.70%	33.78%	41.20%	32.99%	35.87%
Formal, No Shackles	34.18%	35.27%	34.64%	36.54%	29.33%	35.66%
Observations	2,200	13,940	12,260	3,880	1,186	14,880

Notes: Individuals are considered to have education and income similar to a judge if they have a doctoral or professional degree and if their income is \$75,000 per year or more. Individuals are considered to have characteristics similar to seated jurors if they have at least some college education and are employed. Mock jurors who elect not to report their income are not included in either the group of mock jurors with education and income similar to judges or the group of mock jurors with education and income distinct from judges.

At first glance, there are a couple raw differences that are notably large. Mock jurors with prior jury service convict defendants in institutional attire and shackles at higher rates than other defendants, mock jurors without the most common characteristics of seated jurors convict defendants in casual attire less often than other defendants, and mock jurors with education and income levels similar to judges seem to be impacted by shackling in a way that differs based on the attire of the defendant: Shackles increase the odds of conviction for defendants in formal attire and casual attire, and shackles decrease the odds of conviction for defendants in institutional attire. In order to test the significance of these raw differences, Equations 1 and 2 are estimated using linear probability models for the subgroups of mock jurors displayed in Table 6.

Only mock jurors without the common characteristics of seated jurors (i.e., no college education, unemployed, or both) are significantly influenced by the appearance of the defendant. These mock jurors are about 4 percentage points more likely to convict defendants in formal attire than defendants in casual attire. This effect is similar to that experienced by mock jurors with some criminal history, who convict defendants in formal attire significantly more often than defendants in other attire. However, the two effects are not overlapping. Over 80 percent of mock jurors without the most common characteristics of seated jurors have no criminal history.

Overall, the results in this Part indicate that institutional attire can increase the odds of conviction. Men, Republicans, religious mock jurors, and mock jurors with no criminal history convict defendants in institutional clothing significantly more often than defendants in formal attire, and these effects appear to be additive rather than duplicative. Black mock jurors also convict defendants in institutional clothing significantly more often than defendants in formal attire when no shackling is involved, and, although the difference does not reach significance, mock jurors with prior jury service convict defendants in institutional attire and shackles at the highest rate.⁷⁴ Conversely, those with a criminal history convict defendants in formal attire without shackles more often than defendants in casual attire without shackles and defendants in institutional attire with shackles. These competing effects help explain why there are no significant effects of defendant appearance when all mock jurors are pooled together—some of the effects cancel each other out in the aggregate.

The results in this Part also indicate that casual attire can decrease the odds of conviction. Mock jurors between the ages of 25 and 34 convict defendants who are in institutional attire

⁷⁴ Although it is not possible to know for sure, this lack of significance could be the result of a small sample size. Besides the group of mock jurors with education and income levels of judges, the group of mock jurors with prior jury service is the smallest subsample in Table 6.

without shackles significantly more often than defendants in casual attire with shackles, and mock jurors without the common characteristics of seated jurors convict defendants in casual attire significantly less often than defendants in formal attire.

Shackles do not have a clear and consistent effect on conviction decisions. While mock jurors who identify as Democrats are significantly less likely to convict defendants in shackles, mock jurors with education and income levels similar to judges convict defendants in casual or formal attire more often if they are shackled (although the effect is not significant).

C. Differences in Perceptions of the Evidence Do Not Explain the Conviction Results

In Chapter One, biased perceptions of the evidence proved to be a significant mechanism contributing to differences in conviction rates between black and white defendants. That is not the case here. Recall that after rendering a verdict, participants in the first study were asked to rate the believability of each witness and each party's version of events. In most instances, Pearson's chi-squared tests⁷⁵ detect no statistically significant differences in perceptions of the evidence by defendant appearance, indicating that mock jurors are typically not influenced by the defendant's attire or the presence or absence of shackles when assessing the evidence presented at trial. However, in the motive vandalism case scenario, the percent of respondents who find the testimony of the insurance agent and the testimony of the defendant very believable or somewhat believable does differ significantly between the six appearance conditions. Figures 4–7 illustrate participants' assessments of the evidence by defendant appearance for each of the case scenarios.

⁷⁵ In essence, this statistical test determines whether the proportions of respondents finding the relevant aspect of trial somewhat or very believable are equivalent across the six appearance conditions. It is appropriate because both variables being examined (in this case believability and appearance) are categorical, and the observations in each appearance group are independent.

Figure 4. Believability of Eyewitness Vandalism Trial Aspects by Defendant Appearance

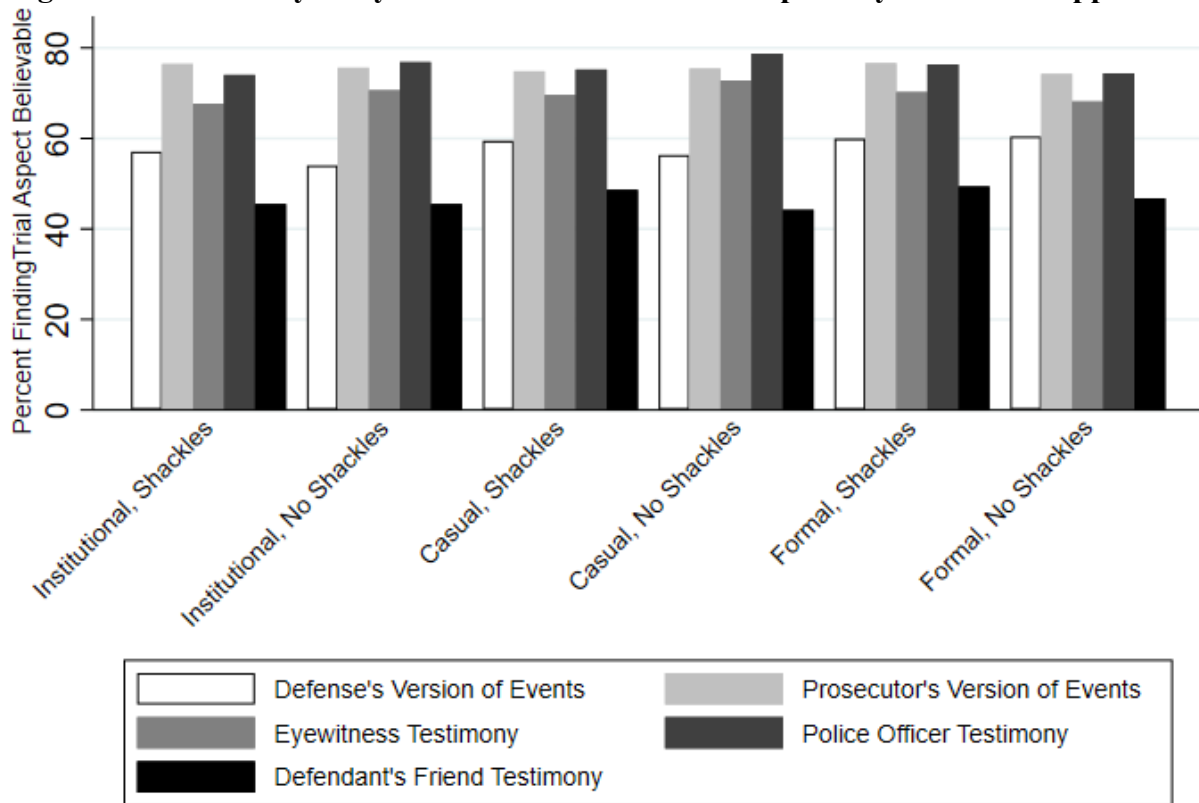


Figure 5. Believability of Eyewitness Arson Trial Aspects by Defendant Appearance

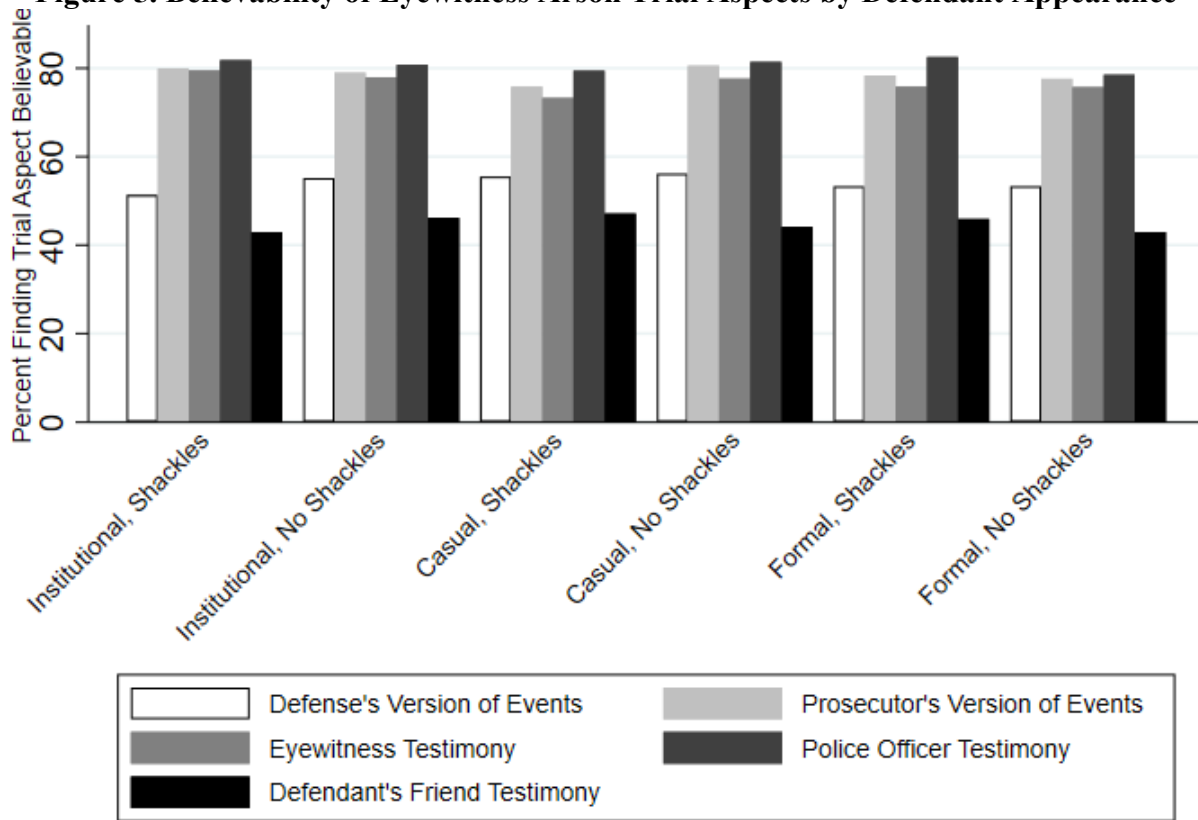


Figure 6. Believability of Motive Vandalism Trial Aspects by Defendant Appearance

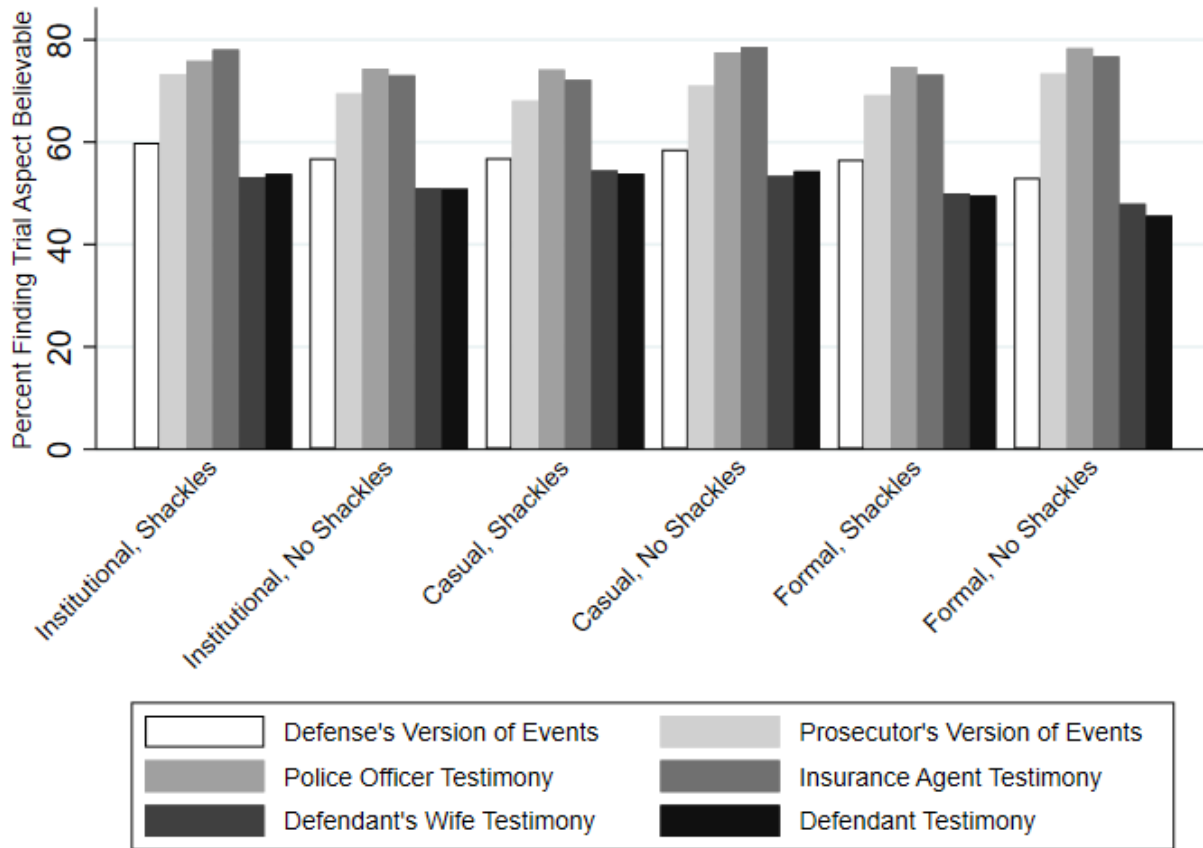
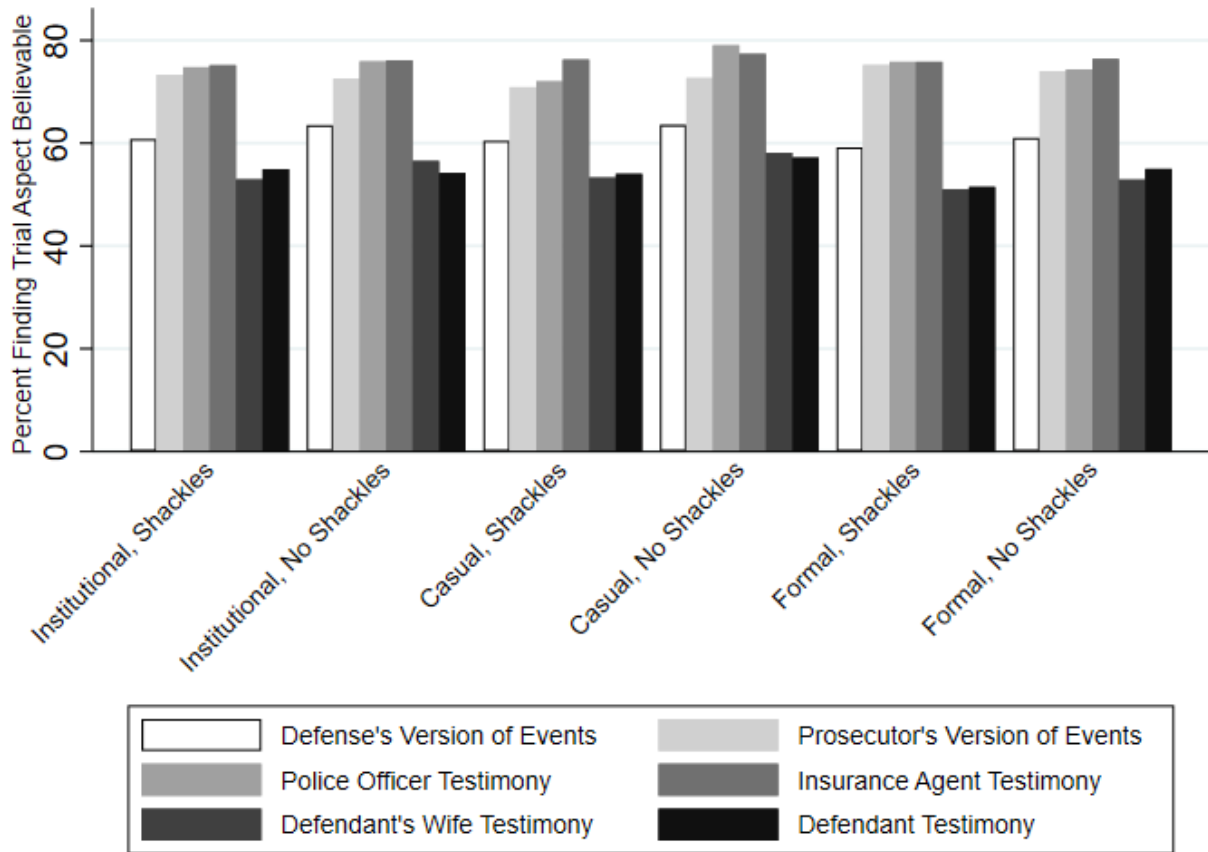


Figure 7. Believability of Motive Arson Trial Aspects by Defendant Appearance



In order to examine what aspects of appearance drive these significant results, Equations 3 and 4 are estimated using linear probability models with data from the motive vandalism case scenario.

$$InsuranceAgentBelievable_i = \alpha + \beta_1 Casual_i + \beta_2 Institutional_i + \beta_3 Shackles_i + \beta_4 Arson_i + \beta_5 Eyewitness_i + \epsilon_i \quad (3)$$

$$DefendantBelievable_i = \alpha + \beta_1 Casual_i + \beta_2 Institutional_i + \beta_3 Shackles_i + \beta_4 Arson_i + \beta_5 Eyewitness_i + \epsilon_i \quad (4)$$

$InsuranceAgentBelievable_i$ and $DefendantBelievable_i$ are dummy variables equal to one if the mock juror found the testimony of the insurance agent and the defendant somewhat or very believable on a five point Likert scale, respectively. $Casual_i$ and $Institutional_i$ are dummy variables equal to one if the defendant is in casual and institutional attire, respectively. Formal

attire is the omitted clothing category. *Shackles_i* is a dummy variable equal to one if the defendant is shackled. *Arson_i* is a dummy variable equal to one when the crime charged is arson, and *Eyewitness_i* is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering on a motive. Standard errors are clustered at the individual level. The results are presented in Table 7.

Table 7. Regression Analyses of Defendant Dress and Shackling on the Believability of the Insurance Agent’s and the Defendant’s Testimony in Motive Vandalism Cases

	Insurance Agent	Defendant
Casual	0.40%	6.52%***
Institutional	0.47%	4.76%**
Shackles	-1.66%	2.13%
Observations	3,989	3,989
R-Squared	0.0004	0.004

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

Notes: Standard ordinary least squares regressions are used. The outcome variable in column one is a dummy variable equal to one if the insurance agent’s testimony was somewhat or very believable on a five point Likert scale. The outcome variable in column two is a dummy variable equal to one if the defendant’s testimony was somewhat or very believable on a five point Likert scale. Casual and Institutional are dummy variables equal to one if the defendant is in that attire. Shackles is a dummy variable equal to one if the defendant is shackled. Standard errors are clustered at the individual level.

Defendants in formal attire are significantly less likely to have their testimony believed than defendants in either institutional or casual attire. This result runs somewhat counter to the results of the previous Part—that defendants in formal attire are typically convicted less often than defendants in institutional attire. In the motive vandalism case scenario, the defendant testifies that he has an insurance policy, he has \$30,000 in credit card debt, he was home with his wife on the night of the incident and didn’t commit the crime, and the empty paint cans found in his home were from an old home improvement project. The only element of this testimony that is different from the motive arson cases is the excuse for the paint cans. Therefore, the negative effect of formal attire on defendant believability must either be based on a perception that defendants who dress up for trial are not the type of people that would engage in home

improvement projects or be crime-specific. For example, mock jurors might think that defendants who dress up for trial even though they are only charged with a minor crime like vandalism are trying to distract or mislead the jury. Whatever the cause of the difference in believability, it is not enough to translate into significant differences in conviction rates. In fact, for many subsamples of mock jurors, another effect of defendant appearance, which will be discussed in the next Part, outweighs this effect such that defendants in formal attire are convicted less often than defendants in institutional attire.

There are no significant effects of defendant dress or shackling alone on the believability of the insurance agent's testimony in the motive vandalism case scenario, indicating that the significant difference detected by the Pearson's chi-squared test must be the result of an interaction between defendant dress and shackling. In order to further explore this result, Equation 5 is estimated using a linear probability model with the data from the motive vandalism cases.

$$InsuranceAgentBelievable_i = \alpha + \beta_1 AppearanceCondition_i + \epsilon_i \quad (5)$$

InsuranceAgentBelievable_i is a dummy variable equal to one if the respondent found the insurance agent's testimony very or somewhat believable. *AppearanceCondition_i* is vector of dummy variables equal to one for five of the six appearance conditions; formal attire with no shackles is the omitted category. Standard errors are clustered at the individual level.

Mock jurors are significantly more likely to find the insurance agent's testimony believable when the defendant is in either casual attire without shackles or institutional attire with shackles than when the defendant is in institutional attire without shackles, casual attire with shackles, or formal attire with shackles. The differences in the probabilities that the insurance agent's testimony will be believed are all around 5–6 percentage points. In the motive

vandalism cases, the insurance agent testifies that the defendant owns the company that rents the filtration system to the pool and that the defendant has an insurance policy on the filtration system. Only the fact that the defendant owns a company that rents pool equipment is unique to this case scenario. It is possible that respondents have a harder time believing that defendants who could not post bail—as evidenced by wearing shackles or institutional clothing—could own a company that rents filtration systems. However, this requires mock jurors to understand and be thinking about the bail system, and it does not explain why no such effect exists when the defendant is in both institutional attire and shackles. While there is no way to know for sure with the data available, it is possible that defendants in both institutional attire and shackles are rated as more believable due to a sympathy effect. Whatever the reason, these differences, like the differences in the believability of the defendant, do not translate into significant differences in conviction rates.

D. Defendants in Institutional Clothing Are Often Presumed Less Innocent

As illustrated by Figure 1, theory provides three possible avenues through which defendant appearance could influence verdict determinations—by affecting the presumption of innocence, by affecting perceptions of the evidence presented at trial, and by affecting the application of the burden of proof. The results from the first study illustrate that defendant appearance has no significant effect on conviction decisions when all mock jurors are pooled together. However, when subgroups of mock jurors are considered independently, they are influenced by defendant dress in two ways: Defendants are often convicted at higher rates when they are in institutional clothing as compared to formal clothing, and mock jurors sometimes convict defendants in casual clothing less often than other defendants. These subgroup effects of defendant dress are not explained by mock jurors' perceptions of the evidence. If anything, mock

jurors' perceptions of the evidence put defendants in formal attire at a disadvantage relative to other defendants. These seemingly conflicting results could be explained by effects of defendant appearance on the presumption of innocence and the burden of proof, which are examined in the second study.

Recall that in the second study participants were asked how likely they thought it was that the defendant was guilty before seeing any evidence. Pooled together, mock jurors' pre-evidence assessments of guilt are not significantly influenced by defendant appearance. Table 8 presents and Figure 8 depicts these assessments by crime charged and defendant appearance. One-way ANOVA tests detect no significant differences in average pre-evidence assessments of guilt between the six defendant appearance variations in any case scenario or with all of the cases scenarios combined.⁷⁶

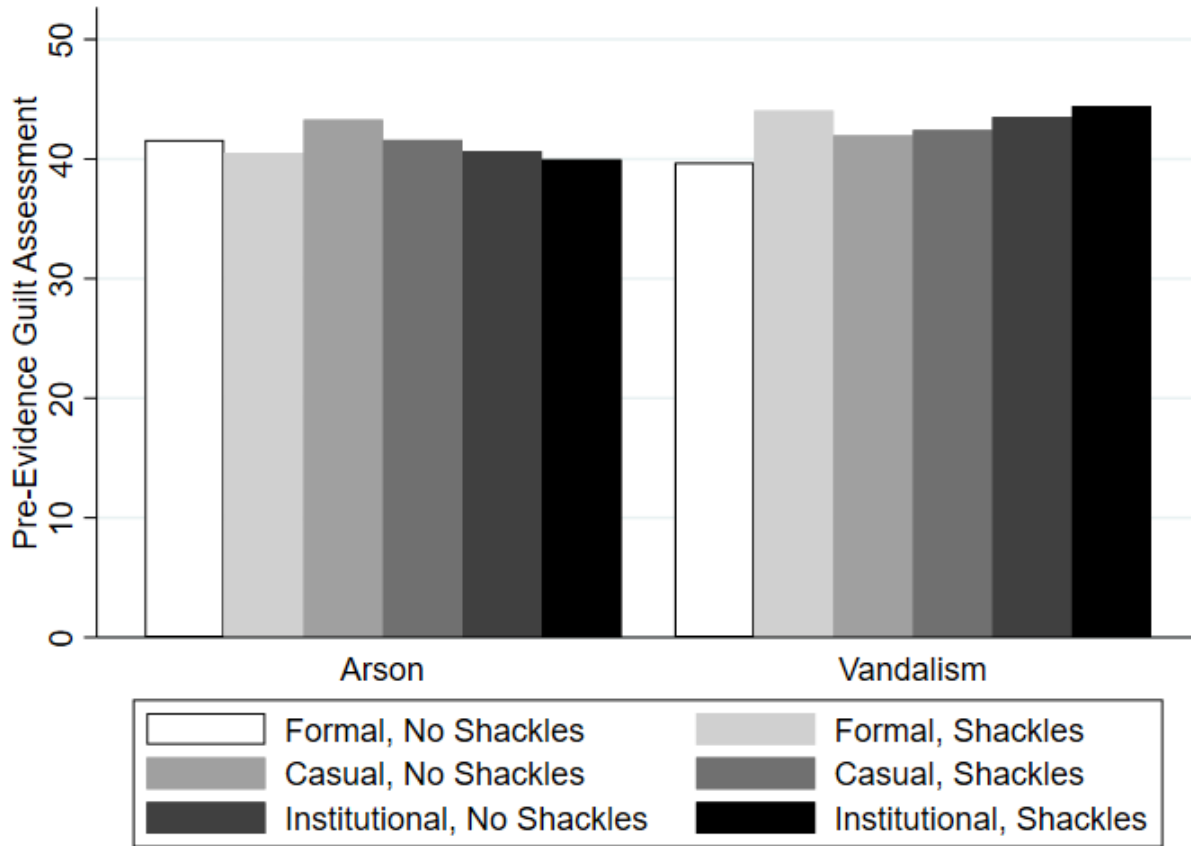
Table 8. Pre-Evidence Assessments of Guilt by Defendant Appearance

	All Cases	Vandalism	Arson
All Defendants	42.01%	42.73%	41.31%
Institutional, Shackles	42.03%	44.44%	39.95%
Institutional, No Shackles	42.11%	43.51%	40.65%
Casual, Shackles	42.02%	42.42%	41.59%
Casual, No Shackles	42.68%	41.97%	43.29%
Formal, Shackles	42.39%	44.05%	40.53%
Formal, No Shackles	40.67%	39.71%	41.59%

Notes: Pre-evidence guilt assessments are the probabilities of guilt assigned to the defendant before any evidence is presented, restricted to between 0 and 100, inclusive. One-way ANOVA tests indicate that pre-evidence guilt assessments do not differ significantly based on defendant appearance (a combination of attire and shackling).

⁷⁶ In essence, this statistical test determines whether average pre-evidence guilt assessments are equivalent across the six appearance conditions. It is appropriate because the dependent variable (in this case pre-evidence assessment of guilt) is continuous, the independent variable (in this case appearance) is categorical, and the observations in each appearance group are independent.

Figure 8. Presumption of Innocence by Defendant Appearance



Even though it does not reach significance, there is a difference of almost 5 percentage points between the presumption of innocence applied to defendants in institutional clothing and shackles and defendants in formal clothing with no shackles in vandalism cases. In order to examine whether this difference helps explain the effects of defendant attire on conviction decisions made by certain subgroups of mock jurors, Equation 6 is estimated using a linear probability model for male mock jurors, Republican mock jurors, religious mock jurors, and mock jurors with no criminal history.

$$PreEvidenceGuiltAssessment_i = \alpha + \beta_1 Casual_i + \beta_2 Institutional_i + \beta_3 Shackles_i + \beta_4 Arson_i + \epsilon_i \quad (6)$$

$PreEvidenceGuiltAssessment_i$ is equal to the probability of guilt that the mock juror assigned to the defendant prior to seeing any evidence. $Casual_i$ and $Institutional_i$ are dummy variables equal to one if the defendant is in casual and institutional attire, respectively. Formal attire is the omitted clothing category. $Shackles_i$ is a dummy variable equal to one if the defendant is shackled. $Arson_i$ is a dummy variable equal to one when the crime charged is arson. In order to maintain adequate sample sizes, vandalism and arson cases are pooled for this analysis. Robust standard errors are used. The results are presented in Table 9.

Table 9. Regression Analyses of Defendant Dress and Shackling on Pre-Evidence Guilt Assessments of Republican, Male, and Religious Mock Jurors and Mock Jurors with No Criminal History

	Republican	Male	Religious	No Criminal History
Casual	1.19%	2.16%	1.80%	0.14%
Institutional	2.98%	4.34%**	2.77%	0.49%
Shackles	1.85%	1.58%	1.83%	0.27%
Arson	-0.39%	-2.09%	-1.91%	-1.88%*
Observations	745	1,144	854	2,007
R-Squared	0.004	0.008	0.005	0.002

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

Notes: Standard ordinary least squares regressions are used. The outcome variable is the probability of guilt assigned to the defendant before any evidence is presented, restricted to between 0 and 100, inclusive. Casual and Institutional are dummy variables equal to one if the defendant is in that attire. Shackles is a dummy variables equal to one if the defendant is shackled. Arson is a dummy variable equal to one if the crime charged is arson. Robust standard errors are used.

Male mock jurors believe that defendants in institutional attire are about 4 percentage points more likely to be guilty than defendants in formal attire before seeing any evidence. While the effect of institutional attire relative to formal attire is not significant for the other subgroups, it is approaching significance for the subgroups of Republican mock jurors and religious mock jurors (with p-values of 0.187 and 0.197, respectively). These groups have a fairly small sample size, which may be responsible for the lack of significance. Overall, these results provide some evidence that at least some of the same subgroups of mock jurors who convict defendants in institutional attire significantly more often than defendants in formal attire also assign defendants

in institutional attire significantly higher pre-evidence assessments of guilt than defendants in formal attire.

These results suggest that the source of any bias against defendants in institutional attire is pre-existing, rather than being formulated based on perceptions of the evidence at trial. These subgroups of mock jurors might perceive defendants in institutional attire as more likely to commit crime generally, more likely to commit vandalism and arson specifically, or they might see institutional attire as a signal of guilt based on their faith in the justice system—in other words, they may believe that only actually guilty defendants will appear at trial in prison garb. In order to explore these possibilities, Table 10 contains the differences in the percent of mock jurors in these subgroups who say that (1) the defendant’s propensity to commit crime is important, (2) the crime charged is important, and (3) their beliefs about the criminal justice system are important in assessing the likelihood that the defendant is guilty pre-evidence when the defendant is in institutional clothing and formal clothing.

Table 10. Relative Importance of the Defendant’s Perceived Propensity to Commit Crime, the Crime Charged, and the Mock Juror’s Beliefs About the Criminal Justice System to the Presumption of Innocence Provided to Defendants in Institutional and Formal Attire by Republican, Male, and Religious Mock Jurors and Mock Jurors with No Criminal History

	Republican	Male	Religious	No Criminal History
Defendant’s Propensity	4.09%	-3.60%	-5.44%	-5.80%**
Crime Charged	0.14%	1.26%	-3.97%	-1.92%
System Beliefs	11.28%**	8.41%**	4.77%	4.41%
Observations	478	758	550	1,293

Notes: Numbers represent the difference in the percent of respondents in the various subgroups who said that the defendant’s propensity to commit crime, the crime charged, and their beliefs about the criminal justice system are somewhat or very important on a five point Likert scale in assessing pre-evidence guilt when the defendant is in institutional and formal clothing. Positive percentages indicate that the factor is more important when the defendant is in institutional attire, and negative percentages indicate that the factor is more important when the defendant is in formal attire. *** p<0.01, ** p<0.05, * p<0.1.

The results suggest that the bias against defendants in institutional attire by Republican and male mock jurors may be the result of faith in the justice system. Indeed, male mock jurors

are about 9 percentage points less likely to say that they distrust the justice system than female mock jurors, and Republican mock jurors are more than 25 percentage points less likely to say that they distrust the justice system than those who identify as a Democrat or independent. The point estimates for religious mock jurors and mock jurors with no criminal history also suggest that they rely on justice system beliefs more when assessing the pre-evidence guilt of defendants in institutional attire. However, the differences are not significant. Both of these subgroups also rely less on the defendant's propensity to commit crime when the defendant is in formal attire, although the difference is only significant for mock jurors with no criminal history. Considered together, the results indicate that the bias exhibited by these subgroups may arise from a combination of faith in the justice system and a belief that defendants in formal attire have a lower propensity to commit crime.

The presumption of innocence results help explain why some subgroups of mock jurors convict defendants in institutional attire more often than defendants in formal attire, but the smaller conviction rates of defendants in casual attire are not explained by either perceptions of the evidence or pre-existing beliefs that influence the presumption of innocence. The next Part will examine whether these differences can be explained by different interpretations of the burden of proof.

E. Defendants in Casual Clothing Receive a Higher Burden of Proof

So far, the conviction results have illustrated that subgroups of mock jurors are influenced by defendant dress in two ways: Defendants are often worse off in terms of conviction rates when they are in institutional clothing as compared to formal clothing, and mock jurors sometimes convict defendants in casual clothing less often than other defendants. The latter of these differences is not explained by differences in the interpretation of the evidence or

differences in the presumption of innocence. This Part explores whether differences in the numerical interpretation of the burden of proof explain this finding.

Recall that in the second study participants were asked how sure they would need to be about the defendant’s guilt in order to convict. In arson cases, one-way ANOVA tests indicate that mock jurors report needing to be significantly more sure in order to convict defendants in casual attire with shackles than defendants in formal attire without shackles.⁷⁷ Table 11 presents and Figure 9 depicts the burden of proof results by defendant appearance. In general, defendants in casual attire benefit from the highest burden of proof interpretations, and defendants in formal attire encounter the difficulty of the lowest burden of proof interpretations. Even in vandalism cases where the effect is insignificant, defendants in casual attire have the first and third highest burden of proof interpretations. Defendants in formal attire with no shackles have the lowest burden of proof interpretation in arson cases by more than 3 percent (although the difference is statistically insignificant), and defendants in formal attire with and without shackles have the lowest burden of proof interpretation in vandalism cases.

Table 11. Burden of Proof Interpretations by Defendant Appearance

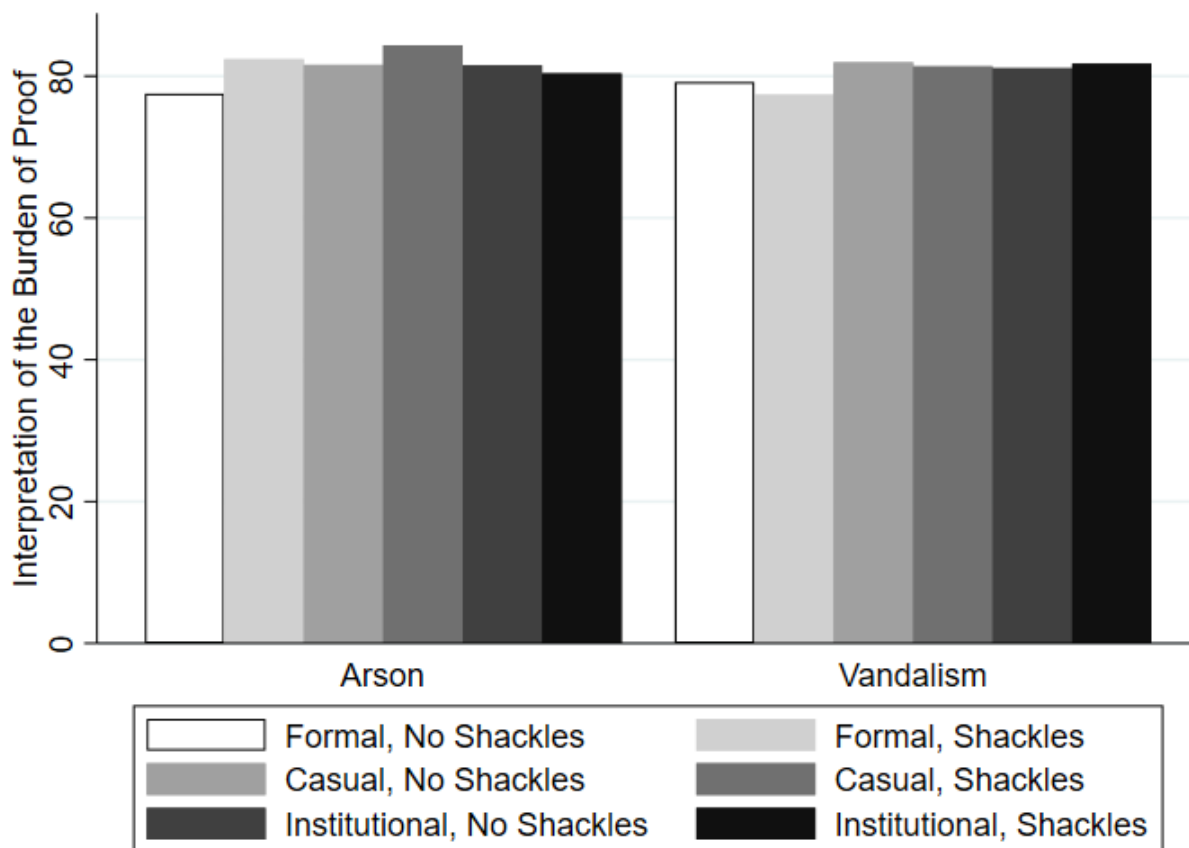
	All Cases	Vandalism	Arson
All Defendants	80.95%	80.53%	81.36%
Institutional, Shackles	81.05%	81.77%	80.41%
Institutional, No Shackles	81.34%	81.14%	81.54%
Casual, Shackles	82.81%	81.40%	84.36% ^a
Casual, No Shackles	81.78%	81.97%	81.62%
Formal, Shackles	79.76%	77.37%	82.42%
Formal, No Shackles	78.35%	79.19%	77.55% ^a

^a different with $p < 0.05$

Notes: Burden of proof interpretations (i.e., how sure jurors reported needing to be about the defendant’s guilt to return a verdict of guilty) are restricted to between 0 and 100, inclusive. One-way ANOVA tests indicate that the difference in the burden of proof applied to defendants in formal clothing with no shackles and defendants in casual clothing with shackles in arson cases is significantly different at the 5 percent level.

⁷⁷ In essence, this statistical test determines whether average numerical interpretation of the burden of proof is equivalent across the six appearance conditions. It is appropriate because the dependent variable (in this case numerical interpretations of the burden of proof) is continuous, the independent variable (in this case appearance) is categorical, and the observations in each appearance group are independent.

Figure 9. Burden of Proof by Defendant Appearance



These results help explain why mock jurors sometimes convict defendants in casual clothing less often than other defendants—they are applying a higher burden of proof, possibly out of a feeling of sympathy for the defendant’s financial situation. As illustrated by Figure 1, interpretations of the burden of proof could be influenced by the perceived cost of a wrongful conviction, the perceived cost of a wrongful acquittal, and the perceived importance of the burden of proof. In order to further investigate the source of the higher burden of proof for defendants in casual attire, Table 12 contains the differences in the percent of mock jurors who said that (1) the cost of a wrongful conviction, (2) the cost of a wrongful acquittal, and (3) whether the prosecutor met the burden of proof would be very or somewhat important on a five

point Likert scale in determining whether to convict when the defendant was in casual clothing and formal clothing.

Table 12. Relative Importance of Wrongful Conviction Costs, Wrongful Acquittal Costs, and the Burden of Proof to the Burden of Proof Applied to Defendants in Casual and Formal Attire

	All Mock Jurors
Wrongful Conviction Cost	2.14%
Wrongful Acquittal Cost	2.92%
Burden of Proof	0.84%
Observations	1,649

Notes: Numbers represent the difference in the percent of respondents who said that the cost of a wrongful conviction, the cost of a wrongful acquittal, and whether the prosecutor met the burden of proof would be somewhat or very important on a five point Likert scale in rendering a verdict when the defendant is in casual and formal clothing. Positive percentages indicate that the factor is more important when the defendant is in casual attire, and negative percentages indicate that the factor is more important when the defendant is in formal attire. *** p<0.01, ** p<0.05, * p<0.1.

None of the differences are significant, indicating that the bias in favor of defendants in casual attire is probably not the result of conscious concern about a higher cost of wrongful convictions, a lower cost of wrongful acquittals, or enforcing the burden of proof. The direction of the coefficients indicate that even if wrongful conviction costs are seen as higher for defendants in casual attire, wrongful acquittal costs are also seen as higher for these defendants. Therefore, the source of the effect of casual attire on the burden of proof is likely either unconscious or unreported.

In summary, when mock jurors are pooled together, there is no significant effect of defendant appearance on conviction decisions, and the effect of defendant dress and shackling does not differ based on the race of the defendant. However, when demographic characteristics are used to isolate subgroups of mock jurors, important effects of defendant dress emerge. Male mock jurors, Republican mock jurors, religious mock jurors, and mock jurors with no criminal history all convict defendants in institutional attire at higher rates than defendants in formal attire, and these results are not duplicative based on overlapping demographics. Black mock

jurors also convict defendants in institutional clothing more often than defendants in formal attire when no shackling is involved. For male and republican mock jurors, the bias against defendants in institutional attire appears to be related to greater faith in the justice system, and for mock jurors with no criminal history, the bias against defendants in institutional attire appears to be related to a belief that defendants in formal attire have a lower propensity to commit crime.

Additionally, some subgroups of mock jurors convict defendants in casual attire less often than other defendants. Mock jurors between the ages of 25 and 34 convict defendants who are in institutional attire without shackles significantly more often than defendants in casual attire with shackles, and mock jurors without the common characteristics of seated jurors convict defendants in casual attire significantly less often than defendants in formal attire. This effect appears to be related to significant differences in the numerical interpretation of the burden of proof—mock jurors see “beyond a reasonable doubt” as a stricter standard in cases where the defendant is wearing casual clothing, at least when the crime charged is arson. However, these differences in the numerical interpretation of the burden of proof are not correlated with differences in reported concerns over the cost of wrongful convictions, the cost of wrongful acquittals, or the burden of proof threshold being met.

Results related to the effect of mock juror criminal history may be of particular interest to policymakers since these individuals are often struck from the jury during voir dire or prohibited from serving on a jury altogether (Kalt 2003; Salyers 1999). Mock jurors with some criminal history convict defendants in formal attire without shackles more often than defendants in institutional attire with shackles and defendants in casual attire without shackles. This result suggests that defendants in formal attire without shackles are relatively better off and defendants

in casual attire without shackles or institutional attire with shackles are relatively worse off in jurisdictions that prevent convicted felons from serving on the jury.

The results based on judge and juror demographics are also of particular importance because they shed light on the external validity of the overarching results. Only mock jurors without the most common characteristics of seated jurors are significantly influenced by defendant appearance (they convict defendants in formal attire more often than defendants in casual attire), but it is likely that subgroups of actual judges and jurors, just like subgroups of all mock jurors, are susceptible to influence by defendant appearance. The next Section will explore whether defendant appearance influences sentencing decisions.

VII. Casual Attire Decreases Sentence Length and Institutional Attire Increases Sentence Length

After rendering a verdict, all mock jurors in the first study were asked to assume the role of a sentencing judge, assume a jury had found the defendant guilty, and select a nonzero sentence for the defendant, regardless of their verdict decision. In vandalism cases, mock jurors rendered a sentence between one and five years, inclusive, and in arson cases, mock jurors rendered a sentence between one and ten years, inclusive. This Section will examine the impact of defendant appearance on assigned sentence length.

The motive scenarios result in greater sentences than the eyewitness scenarios, and, unsurprisingly, arson cases result in longer sentences than vandalism cases.⁷⁸ Figures 10 and 11 display the distribution of sentences in each of the four case scenarios. As discussed in Chapter One, the distribution is left-skewed in all of the case scenarios, indicating a preference for lighter

⁷⁸ Some of this difference could be due to an anchoring effect because the maximum sentence was higher for arson than for vandalism (Enough & Mussweiler 2001). Such an anchoring effect is not unique to the experimental vignette study context, however, as judges are also exposed to minimum and maximum sentences imposed by Congress and sentence recommendations presented by the defense and prosecution.

sentences. In the eyewitness vandalism case scenario, around 70 percent of respondents assign the minimum sentence of one year. In the motive vandalism case scenario, just over 60 percent of respondents assign the minimum sentence of one year. Among mock jurors who found the defendant guilty, just over 60 percent return the minimum sentence in the eyewitness vandalism case scenario, and about 45 percent return the minimum sentence in the motive vandalism case scenario.

There is significantly more variation in the sentences assigned in arson cases. In the eyewitness arson case scenario, just under 25 percent of respondents assign the minimum sentence of a year, and the median sentence is two years. In the motive arson case scenario, around 16 percent of respondents assign the minimum sentence of a year, and the median sentence is 2.9 years. Of the jurors that return a verdict of guilty, about 14 percent assign the minimum sentence in the eyewitness arson case scenario, and about 5 percent assign the minimum sentence in the motive arson case scenario. About 68 percent of participants feel somewhat confident or very confident in their sentencing decision.

Figure 10. Distribution of Sentences Assigned in Vandalism Cases

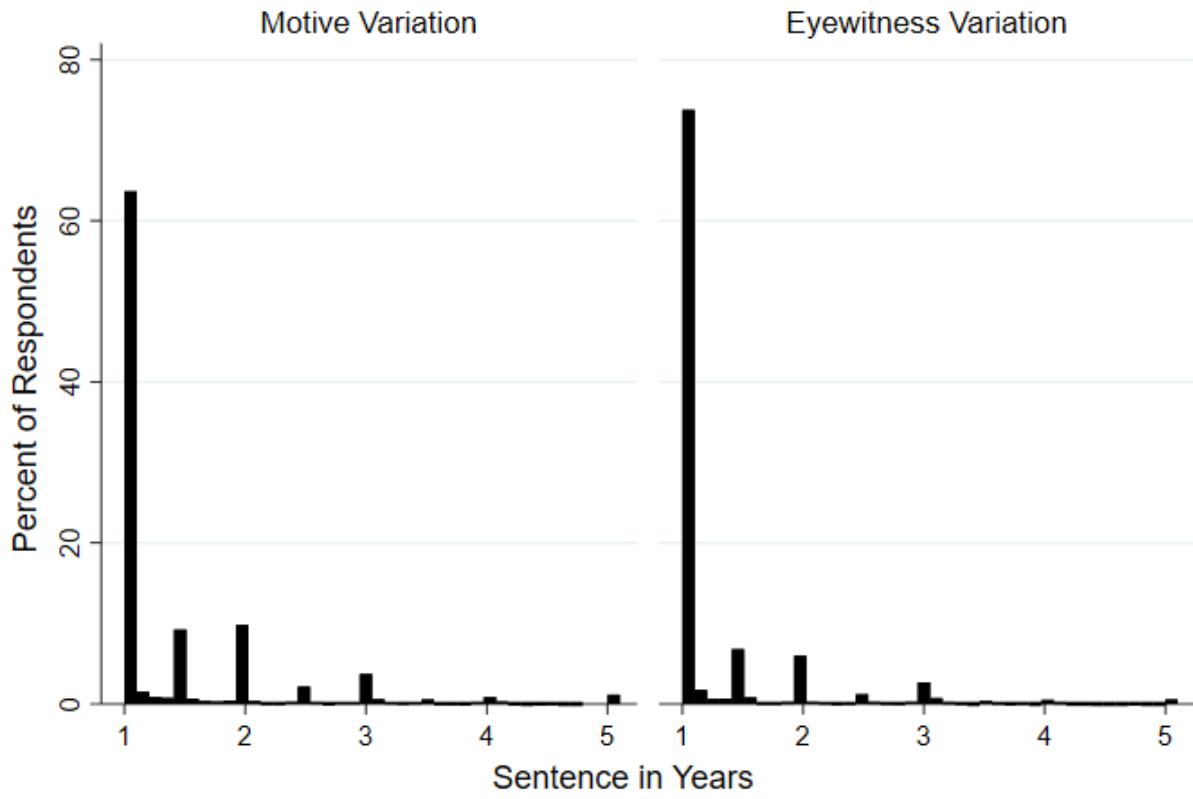
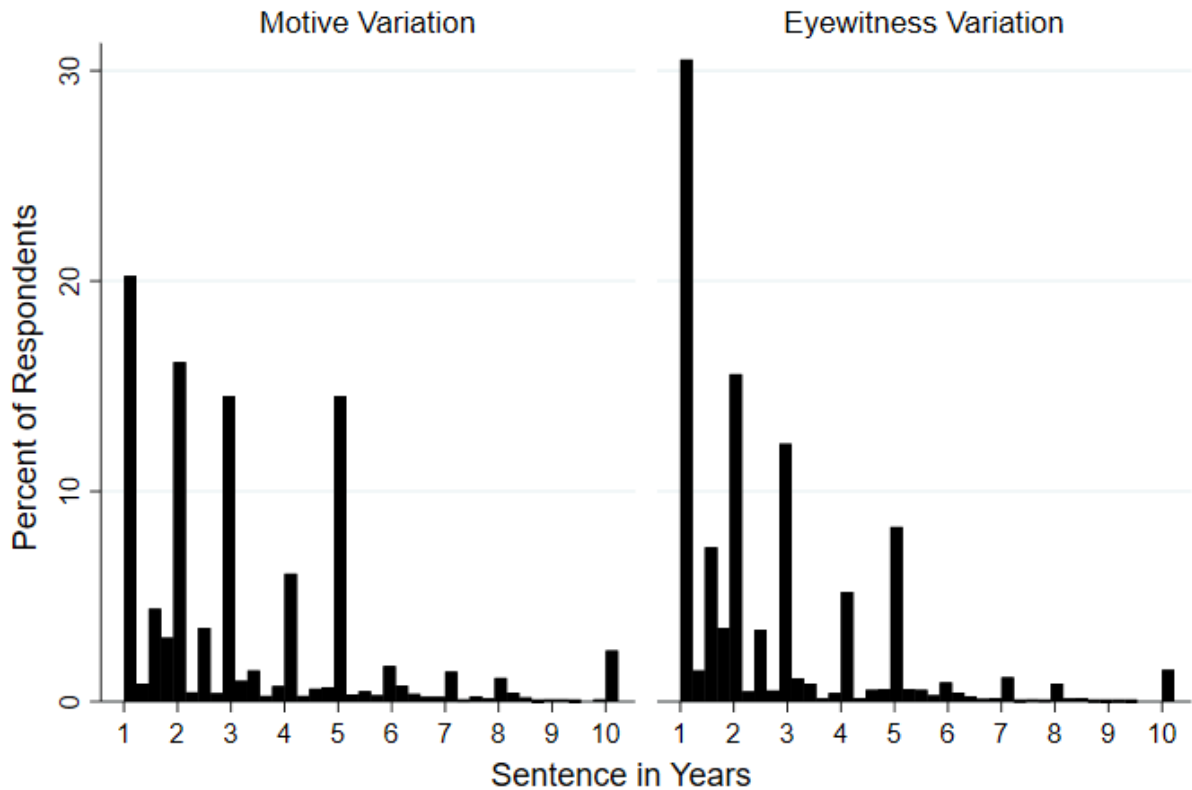


Figure 11. Distribution of Sentences Assigned in Arson Cases



Despite the large number of respondents who return the minimum sentence, there is a significant effect of defendant dress on sentence length when all of the case scenarios are pooled together: Defendants in casual attire receive significantly lighter sentences (by almost 45 days in prison) than defendants in institutional attire. Table 13 contains the average sentences assigned to defendants overall and in each of the case scenarios by appearance condition. T-tests were conducted to determine if the average sentence differs significantly based on the presence or absence of shackles alone, one-way ANOVA tests were conducted to determine if sentences differ significantly based on defendant attire alone, and one-way ANOVA tests were also

conducted to determine if sentences differ significantly between any of the six appearance conditions, allowing for shackling and attire to interact.⁷⁹

Table 13. Sentence Length by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	2.14	1.31	1.45	2.62	3.19
Institutional, Shackles	2.20 ^a	1.32	1.48	2.72	3.20
Institutional, No Shackles	2.20 ^b	1.36	1.47	2.61	3.36
Casual, Shackles	2.13	1.29	1.42	2.67	3.18
Casual, No Shackles	2.04 ^{ab}	1.30	1.39	2.49	3.04
Formal, Shackles	2.14	1.28	1.44	2.56	3.25
Formal, No Shackles	2.15	1.32	1.47	2.69	3.12

^a different with $p < 0.01$; ^b different with $p < 0.01$

Notes: In vandalism cases, sentences are restricted to between 1 and 5 years, inclusive, and in arson cases sentences are restricted to between 1 and 10 years, inclusive. One-way ANOVA tests indicate that the differences in sentences given to defendants in casual clothing with no shackles and defendants in formal clothing (with and without shackles) are significantly different at the 1 percent level when all case scenarios are pooled.

The average sentences assigned to defendants in formal attire lie between those assigned to defendants in casual attire and those assigned to defendants in institutional attire, indicating that there is a benefit to being in casual clothing relative to formal clothing as well as a penalty for being in institutional clothing relative to formal clothing. When the effects of clothing and shackles are allowed to interact, the casual clothing effect becomes more specific—defendants in casual attire without shackles receive significantly lighter sentences (by almost 60 days in prison) than defendants in institutional attire (with or without shackles).

The effect of defendant dress on sentence length is primarily concentrated among white defendants. Tables 14 and 15 contain the average sentences assigned to white and black defendants by case scenario and defendant appearance. The same t-tests and one-way ANOVA tests described above were run for white defendants and black defendants.

⁷⁹ In essence, one-way ANOVA tests determine whether average sentences are equivalent across the three attire conditions and the six appearance conditions. They are appropriate because the dependent variable (in this case sentence) is continuous, the independent variables (in this case appearance and attire) are categorical, and the observations in each appearance group are independent.

Table 14. Sentencing of White Defendants by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	2.19	1.30	1.48	2.69	3.32
Institutional, Shackles	2.32 ^a	1.28	1.52	2.92 ^c	3.36
Institutional, No Shackles	2.28 ^b	1.33	1.50	2.79	3.59
Casual, Shackles	2.16	1.33	1.43	2.65	3.28
Casual, No Shackles	2.08 ^{ab}	1.30	1.45	2.48 ^c	3.13
Formal, Shackles	2.15	1.25	1.42	2.57	3.40
Formal, No Shackles	2.16	1.32	1.48	2.78	3.18

^a different with $p < 0.01$; ^b different with $p < 0.05$; ^c different with $p < 0.05$

Notes: In vandalism cases, sentences are restricted to between 1 and 5 years, inclusive, and in arson cases sentences are restricted to between 1 and 10 years, inclusive. One-way ANOVA tests indicate that the difference in sentences given to white defendants in casual clothing with no shackles and white defendants in formal clothing with shackles is significant at the 1 percent level, and the difference in sentences given to white defendants in casual clothing with no shackles and white defendants in formal clothing without shackles is significant at the 5 percent level when all case scenarios are pooled. One-way ANOVA tests also indicate that the difference in sentences given to white defendants in casual clothing with no shackles and white defendants in formal clothing with shackles is significant at the 5 percent level in eyewitness arson cases.

Table 15. Sentencing of Black Defendants by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	2.10	1.32	1.43	2.56	3.06
Institutional, Shackles	2.09	1.35	1.44	2.53	3.02
Institutional, No Shackles	2.11	1.39	1.45	2.46	3.14
Casual, Shackles	2.10	1.26	1.41	2.68	3.08
Casual, No Shackles	1.99	1.29	1.33	2.51	2.94
Formal, Shackles	2.13	1.32	1.45	2.56	3.12
Formal, No Shackles	2.14	1.32	1.47	2.62	3.06

Notes: In vandalism cases, sentences are restricted to between 1 and 5 years, inclusive, and in arson cases sentences are restricted to between 1 and 10 years, inclusive. One-way ANOVA tests indicate that the sentences assigned to black defendants do not differ significantly based on defendant appearance.

When the effects of defendant dress are considered alone, white defendants are sentenced significantly more harshly when they are in institutional clothing than casual clothing in the eyewitness arson case scenario (a difference of about 110 days). Additionally, white defendants are sentenced significantly more harshly when they are in institutional clothing than casual or

formal attire when all of the cases are pooled together: White defendants in casual attire receive an average sentence about 66 days shorter than white defendants in institutional attire, and white defendants in formal attire receive an average sentence about 52 days shorter than white defendants in institutional attire. When the effects of dress and shackling are allowed to interact, only the differences in sentences given to white defendants in casual attire without shackles and white defendants in institutional attire (with or without shackles) remain significant.

When the effects of defendant dress are considered alone, black defendants are also sentenced significantly more harshly when they are in institutional clothing than casual clothing in the eyewitness vandalism case scenario (by about 37 days). However, the effect is only present in that one case scenario, and the effect goes away when the effects of clothing and shackling are allowed to interact. These results illustrate that white defendants receive the shortest sentences when they are in casual clothing and the longest sentences when they are in institutional clothing, and black defendants' sentences are rarely influenced by their attire.

A. The Sentencing Results are Robust Among Important Subgroups of Mock Jurors

Recall that after rendering a verdict, all mock jurors in the first study were asked to assume the role of a sentencing judge, assume a jury had found the defendant guilty, and select a nonzero sentence for the defendant, regardless of their verdict decision. As strange as it may seem to sentence a defendant after finding him not guilty, this is a situation that many real-world judges may face. In a case where a defendant is found guilty by a jury, the judge might think that, had the trial been a bench trial, he or she would have acquitted. Nevertheless, if the verdict is reasonable, the judge will need to accept the verdict and assign a sentence.

It is possible, however, that judges assign sentences differently in cases where they believe that the defendant is guilty and cases where they would have acquitted the defendant. In

the context of this study, this would equate to defendant appearance having a different effect on the assigned sentence depending on whether the mock juror chose to convict.

In order to examine this possibility, Equations 7 and 8 are estimated for the subsamples of mock jurors that found the defendant guilty and not guilty using linear regression models.

$$Sentence_i = \alpha + \beta_1 Casual_i + \beta_2 Institutional_i + \beta_3 Shackles_i + \beta_4 Arson_i + \beta_5 Eyewitness_i + \epsilon_i \quad (7)$$

$$Sentence_i = \alpha + \beta_1 AppearanceCondition_i + \beta_2 Arson_i + \beta_3 Eyewitness_i + \epsilon_i \quad (8)$$

$Sentence_i$ is equal to the sentence assigned. $Casual_i$ and $Institutional_i$ are dummy variables equal to one if the defendant is in casual and institutional attire, respectively. Formal attire is the omitted clothing category. $Shackles_i$ is a dummy variable equal to one if the defendant is shackled. $AppearanceCondition_i$ is vector of dummy variables equal to one for five of the six appearance conditions; formal attire with no shackles is the omitted category. $Arson_i$ is a dummy variable equal to one when the crime charged is arson, and $Eyewitness_i$ is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering on a motive. In order to maintain adequate sample sizes, all four case scenarios are pooled for this analysis. Standard errors are clustered at the individual level.

The effects of defendant appearance on sentencing are entirely concentrated among mock jurors who found the defendant guilty. When sentenced by these mock jurors, the Equation 7 results indicate that defendants are significantly worse off in institutional attire as compared to both formal and casual attire. Defendants in institutional attire receive a sentence about 66 days longer than defendants in casual attire and about 40 days longer than defendants in formal attire. The Equation 8 results, which allow for defendant dress and shackling to interact, illustrate that defendants in casual attire without shackles are better off when sentenced by mock jurors who convicted. In fact, they receive sentences that are significantly shorter than defendants in all

other appearance conditions. Conversely, mock jurors who acquitted the defendant are not significantly affected by defendant dress or shackling when assigning sentences.

It is also important to ensure that the results do not differ for the subgroups of mock jurors that have characteristics similar to judges and seated jurors. Table 16 contains the average sentences assigned by mock jurors with and without the education and income levels of typical judges, mock jurors with and without the most common characteristics of seated jurors, and individuals with and without prior jury service experience. As before, based on data from the Bureau of Labor Statistics' Occupational Employment Statistics survey, individuals are considered to have the education and income levels of typical judges if they have a doctoral or professional degree and an income of \$75,000 per year or more (Bureau of Labor Statistics 2018). Based on data from a survey administered to seated jurors in four courts by Hannaford-Agor et al. (2003) and the demographic characteristics of participants who report prior jury service in the first study, individuals are considered to have the most common characteristics of seated jurors if they have at least some college education and are employed. In order to maintain adequate sample sizes, all four case scenarios are pooled.

Table 16. Sentence Length by Defendant Appearance, Judicial and Juror Demographics, and Prior Jury Service

	Prior Jury Service	No Prior Jury Service	Characteristics Similar to Seated Jurors	Characteristics Distinct from Seated Jurors	Education and Income Similar to Judges	Education and Income Distinct from Judges
All Defendants	2.15	2.14	2.14	2.14	2.17	2.14
Institutional, Shackles	2.44	2.17	2.10	2.21	2.43	2.19
Institutional, No Shackles	2.15	2.20	2.21	2.14	2.16	2.20
Casual, Shackles	2.06	2.15	2.15	2.16	2.13	2.13
Casual, No Shackles	1.98	2.04	2.06	1.96	2.20	2.02

Formal, Shackles	2.16	2.14	2.12	2.21	1.99	2.15
Formal, No Shackles	2.12	2.16	2.15	2.17	2.12	2.16
Observations	2,200	13,940	12,260	3,880	1,186	14,880

Notes: Individuals are considered to have education and income similar to a judge if they have a doctoral or professional degree and if their income is \$75,000 per year or more. Individuals are considered to have characteristics similar to seated jurors if they have at least some college education and are employed. Mock jurors who elect not to report their income are not included in either the group of mock jurors with education and income similar to judges or the group of mock jurors with education and income distinct from judges.

At first glance, there are a couple raw differences that are notably large. Mock jurors with prior jury service and mock jurors with the income and education levels of typical judges appear to sentence defendants in institutional attire and shackles more harshly than other defendants, and all of the groups except mock jurors with the income and education levels of typical judges appear to sentence defendants in casual attire without shackles more leniently than other defendants. In order to test the significance of these raw differences, Equations 7 and 8 are estimated for the subgroups of mock jurors displayed in Table 16.

All of the subgroups except mock jurors with the education and income levels of typical judges are affected by defendant appearance in some way. The results from Equation 7 illustrate that all of the subgroups except those with the income and education levels of typical judges and those without the most common characteristics of seated jurors sentence defendants in institutional attire to significantly longer sentences than defendants in casual attire. Additionally, mock jurors with prior jury service sentence defendants in formal attire to longer sentences than defendants in casual attire, and mock jurors with the most common characteristics of seated jurors sentence defendants in institutional clothing to longer sentences than defendants in formal clothing. Together, these results indicate that real seated jurors may be more likely to be affected by defendant dress than mock jurors as a whole when they are asked to assign a noncapital sentence.

The results from Equation 8, which allow the effects of dress and shackling to interact, tell a similar story. Mock jurors with prior jury service sentence defendants in institutional clothing with shackles particularly harshly. Mock jurors with no prior jury service sentence defendants in casual clothing without shackles particularly leniently. Mock jurors without the education and income levels of typical judges sentence defendants in casual attire without shackles significantly more leniently than defendants in every other appearance group. Finally, despite not being influenced by defendant dress or shackling alone, mock jurors without the most common characteristics of seated jurors sentence defendants in formal attire and shackles and defendants in institutional attire and shackles significantly more harshly than defendants in casual attire without shackles.

The lack of significant results for mock jurors with the education and income levels of typical judges might indicate that judges are better at tuning out defendant dress when assigning sentences. However, the sample of mock jurors with the education and income levels of typical judges is relatively small, and the raw differences in mean sentences indicate these mock jurors at least have an insignificant tendency to favor defendants in casual attire and disfavor defendants in institutional attire, like all other mock jurors. Therefore, the lack of significant effects for mock jurors with the education and income levels of typical judges should be interpreted with caution.

B. Perceived Recidivism Is Not the Mechanism Through Which Clothing Affects Sentencing

The results presented above suggest that, particularly when they are white, defendants in casual clothing receive the lightest sentences and defendants in institutional attire receive the harshest sentences from mock jurors who convict. Because decisionmakers are typically given

more discretion when assigning a sentence than when reaching a verdict, identifying the channel through which this clothing effect occurs requires sifting through many possible explanations. One such explanation is recidivism predictions. Decisionmakers might believe that defendants in institutional attire are more likely to reoffend and defendants in casual attire are more likely to be one-time offenders.

Because perceived recidivism is one obvious mechanism through which appearance bias might work, exploring this relationship further is warranted. Recall that after sentencing the defendant in the first study, participants were asked to rate how likely they thought it was that the defendant would commit a crime in the next five years if acquitted. Pearson’s chi-squared tests indicate that these recidivism predictions do not differ significantly by defendant appearance in any of the case scenarios.⁸⁰ Table 17 summarizes the predictions by case scenario and defendant appearance.

Table 17. Recidivism Predictions by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	25.86%	29.43%	17.57%	35.00%	21.44%
Institutional, Shackles	27.37%	30.84%	19.91%	35.81%	22.67%
Institutional, No Shackles	26.22%	32.14%	15.16%	36.03%	21.74%
Casual, Shackles	26.06%	29.52%	17.10%	33.07%	24.16%
Casual, No Shackles	26.31%	30.52%	17.45%	34.70%	22.15%
Formal, Shackles	25.26%	25.60%	17.94%	36.88%	20.18%
Formal, No Shackles	23.97%	28.06%	18.02%	32.61%	17.96%

Notes: Numbers represent the percent of respondents who predict that the defendant is very likely or somewhat likely to recidivate in the next five years if acquitted on a five point Likert scale. Pearson’s chi-squared tests indicate that these percentages do not differ significantly by the defendant’s appearance overall or in any of the case scenarios.

⁸⁰ In essence, this statistical test determines whether the proportions of respondents predicting that the defendant is very likely or somewhat likely to recidivate are equivalent across the six appearance conditions. It is appropriate because both variables being examined (in this case recidivism and appearance) are categorical, and the observations in each appearance group are independent.

The effect of appearance on sentencing differs based on the race of the defendant:

Usually, only white defendants' sentences are significantly affected by appearance. Nevertheless, even when white defendants are considered separately from black defendants, recidivism predictions do not differ significantly based on the appearance of the defendant. Tables 18 and 19 contain recidivism predictions by case scenario for white and black defendants, respectively.

Table 18. Recidivism Predictions for White Defendants by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	28.41%	31.35%	18.61%	40.60%	22.94%
Institutional, Shackles	30.60%	33.33%	23.13%	41.35%	24.30%
Institutional, No Shackles	27.13%	32.34%	13.65%	42.41%	20.44%
Casual, Shackles	29.19%	32.18%	19.06%	38.56%	27.54%
Casual, No Shackles	29.64%	33.16%	18.58%	40.70%	25.14%
Formal, Shackles	26.70%	26.42%	18.84%	40.30%	19.81%
Formal, No Shackles	27.08%	30.79%	18.86%	40.20%	19.94%

Notes: Numbers represent the percent of respondents who predict that a white defendant is very likely or somewhat likely to recidivate in the next five years if acquitted on a five point Likert scale. Pearson's chi-squared tests indicate that these percentages do not differ significantly by the defendant's appearance overall or in any of the case scenarios.

Table 19. Recidivism Predictions for Black Defendants by Defendant Appearance

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
All Defendants	23.33%	27.39%	16.59%	29.59%	19.95%
Institutional, Shackles	24.19%	28.40%	16.96%	30.35%	20.94%
Institutional, No Shackles	25.34%	31.94%	16.62%	30.25%	23.01%
Casual, Shackles	22.83%	26.82%	15.00%	28.75%	20.71%
Casual, No Shackles	22.77%	27.67%	16.37%	28.12%	18.85%
Formal, Shackles	23.92%	24.69%	17.22%	33.62%	20.50%
Formal, No Shackles	20.98%	25.00%	17.23%	26.21%	16.09%

Notes: Numbers represent the percent of respondents who predict that a black defendant is very likely or somewhat likely to recidivate in the next five years if acquitted on a five point Likert scale. Pearson's chi-squared tests indicate that these percentages do not differ significantly by the defendant's appearance overall or in any of the case scenarios.

These results suggest that mock jurors' recidivism predictions are not influenced by defendant dress or shackling. This means that white defendants in casual attire receive shorter sentences and white defendants in institutional attire receive longer sentences than other white defendants from mock jurors who convicted despite being perceived as equally likely to recidivate if acquitted. In order to verify this result, Equations 9 and 10 are estimated for each of the four criminal case scenarios and for all of the case scenarios combined using only the sample of cases in which a white defendant was convicted.

$$Sentence_i = \alpha + \beta_1 Casual_i + \beta_2 Institutional_i + \beta_3 Shackles_i + \beta_4 Recidivist_i + \beta_5 Arson_i + \beta_6 Eyewitness_i + \epsilon_i \quad (9)$$

$$Sentence_i = \alpha + \beta_1 AppearanceCondition_i + \beta_2 Recidivist_i + \beta_3 Arson_i + \beta_4 Eyewitness_i + \epsilon_i \quad (10)$$

$Sentence_i$ is equal to the assigned sentence length. $Casual_i$ and $Institutional_i$ are dummy variables equal to one if the defendant is in casual and institutional attire, respectively. Formal attire is the omitted clothing category. $Shackles_i$ is a dummy variable equal to one if the defendant is shackled. $AppearanceCondition_i$ is a vector of dummy variables equal to one for five of the six appearance conditions; formal attire with no shackles is the omitted category. $Recidivist_i$ is a dummy variable equal to one if the respondent said they thought it was very likely or somewhat likely that the defendant would recidivate in the next five years if acquitted. $Arson_i$ is a dummy variable equal to one when the crime charged is arson, and $Eyewitness_i$ is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering on a motive. These last two control variables are only included when all of the case scenarios are combined. Standard errors are clustered at the individual level. The Equation 9 results are presented in Table 20.

Table 20. Regression Analyses of White Defendant Dress, Shackling, and Recidivism Predictions on Sentencing by Mock Jurors Who Convicted

	All Case Scenarios	Eyewitness Vandalism	Motive Vandalism	Eyewitness Arson	Motive Arson
Casual	-0.11*	0.05	-0.08	-0.27	-0.17
Institutional	0.20***	0.09	0.08	0.37*	0.23
Shackles	0.05	-0.03	-0.03	0.18	0.05
Recidivist	0.51***	0.18***	0.49***	0.74***	0.63***
Arson	1.79***	—	—	—	—
Eyewitness	-0.71***	—	—	—	—
Observations	3,141	772	610	754	1,005
R-squared	0.273	0.014	0.062	0.044	0.034

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

Notes: Standard ordinary least squares regressions are used. The outcome variable is the sentence assigned, between one and five years, inclusive, for vandalism and between one and ten years, inclusive, for arson. Casual and Institutional are dummy variables equal to one if the defendant is in that attire. Shackles is a dummy variable equal to one if the defendant is shackled. Recidivist is a dummy variable equal to one if the respondent predicted that the defendant is somewhat or very likely to recidivate in the next five years if acquitted on a five point Likert scale. Arson is a dummy variable equal to one if the crime charged is arson. Eyewitness is a dummy variable equal to one if the evidence revolves around eyewitness testimony rather than a motive. Standard errors are clustered at the individual level.

Even after accounting for recidivism predictions, white defendants in casual attire receive the lightest sentences, and white defendants in institutional attire receive the harshest sentences from mock jurors who convicted. In motive vandalism cases, eyewitness arson cases, motive arson cases, and when all of the case scenarios are pooled, white defendants in casual attire receive significantly shorter sentences than white defendants in institutional attire (by 58, 234, 146, and 113 days, respectively). Additionally, when all of the case scenarios are pooled, white defendants in casual attire receive significantly shorter sentences than white defendants in formal attire (by 40 days), and in eyewitness arson cases and when all of the case scenarios are pooled, white defendants in formal attire receive significantly shorter sentences than white defendants in institutional attire (by 135 and 73 days, respectively).

The Equation 10 results, which allow the effects of defendant dress and shackles to interact, tell a similar story. When all of the case scenarios are pooled, white defendants in casual attire without shackles are sentenced particularly leniently and white defendants in institutional

attire are sentenced particularly harshly by mock jurors who convicted. In the eyewitness arson case scenario, white defendants in casual attire without shackles receive significantly shorter sentences than white defendants in institutional attire (with or without shackles) and white defendants in formal attire without shackles, and white defendants in institutional attire with shackles receive significantly longer sentences than white defendants in formal attire with shackles and white defendants in casual attire with shackles. In motive arson cases, white defendants in casual attire without shackles receive significantly shorter sentences than white defendants in institutional attire (with or without shackles) and white defendants in formal attire with shackles, and white defendants in institutional attire (with or without shackles) receive significantly longer sentences than white defendants in casual attire with shackles and white defendants in formal attire without shackles.

Again, these results suggest that white defendants in casual attire (specifically, without shackles) and white defendants in institutional attire are not receiving significantly different sentences from mock jurors who convicted because they are perceived as more or less likely to reoffend. Rather, they are receiving significantly different sentences *despite* being perceived as equally likely to reoffend. If decisionmakers are sentencing defendants based on the goals of incapacitation, deterrence, retribution, and rehabilitation, and perceived recidivism is not the mechanism, then these mock jurors must view white defendants in casual attire without shackles as less culpable, more capable of rehabilitation, or less likely to provide general deterrence as compared to white defendants in institutional attire. However, this study is not able to tease out these potential explanations, so discerning the precise source of this effect remains an important area for future research.

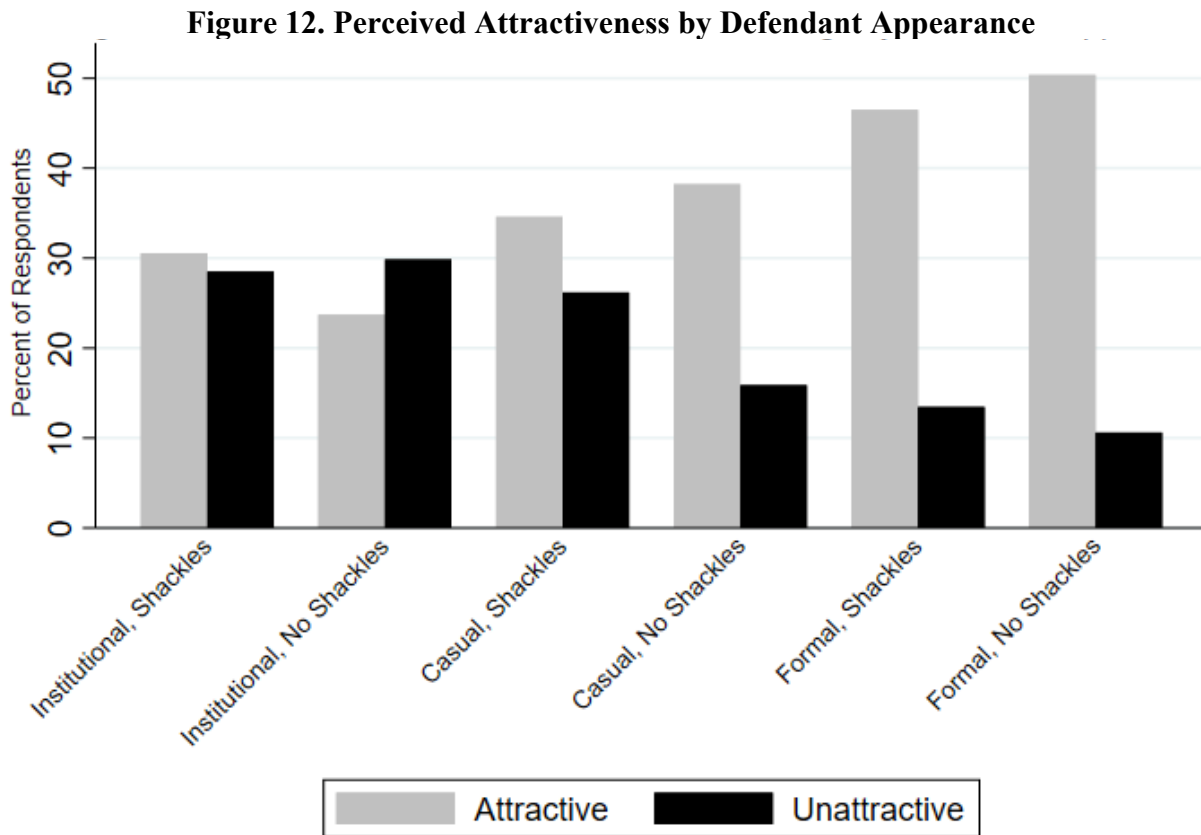
Importantly, the sentencing results also indicate that black defendants' sentences are only rarely significantly influenced by appearance, mock jurors who acquitted the defendant are not influenced by defendant appearance when assigning sentences, and mock jurors with the education and income levels of typical judges—who are typically responsible for assigning sentences in noncapital cases—may be better at tuning out defendant appearance when assigning sentences. However, the sample of mock jurors with the education and income levels of typical judges is relatively small, and the raw differences in mean sentences indicate these mock jurors at least have an insignificant tendency to favor defendants in casual attire and disfavor defendants in institutional attire, like all other mock jurors.

C. Perceived Attractiveness Is Not the Mechanism Through Which Clothing Affects Sentencing

Attractiveness is a particularly complex aspect of appearance because it interacts with so many other aspects of appearance. A defendant wearing a suit may be viewed as less dangerous than a defendant wearing jeans and a t-shirt, but it is possible that the defendant in jeans and a t-shirt is viewed as more attractive than the defendant in a suit. Since previous studies have found that defendant attractiveness impacts guilt (Abwender & Hough 2001; Desantis & Kayson 1997; Efran 1974; Leventhal & Kreate 1977; MacCoun 1990; Patry 2008) and sentencing (Abwender & Hough 2001; Efran 1974; Friend & Vinson 1974; Leventhal & Kreate 1977; Piehl 1977; Sigall & Ostrove 1975; Smith & Hed 1997) decisions, it is important to examine how the manipulations of defendant dress and shackling in this study relate to attractiveness.

Appearance and attractiveness are highly correlated such that defendants in formal attire (with and without shackles) and defendants in casual attire without shackles are seen as the most attractive, and defendants in the other appearance conditions are seen as the least attractive.

Recall that second study respondents were asked to rate the attractiveness of one of the twenty-four defendant pictures on a five point Likert scale. Figure 12 illustrates these attractiveness assessments by appearance condition. A respondent is coded as finding a defendant unattractive if the respondent said the defendant was very or somewhat unattractive and as finding a defendant attractive if the respondent said the defendant was very or somewhat attractive. Pearson’s chi-squared tests confirm that there are significant differences in both attractiveness and unattractiveness by appearance condition.



In order to investigate whether the significant effects of appearance on attractiveness explain the significant differences in sentencing by appearance, Equations 11-14 are estimated using data from the first study.

$$Sentence_i = \alpha + \beta_1 Casual_i + \beta_2 Institutional_i + \beta_3 Shackles_i + \beta_4 Attractive_i + \beta_5 Arson_i + \beta_6 Eyewitness_i + \epsilon_i \quad (11)$$

$$Sentence_i = \alpha + \beta_1 Casual_i + \beta_2 Institutional_i + \beta_3 Shackles_i + \beta_4 Unattractive_i + \beta_5 Arson_i + \beta_6 Eyewitness_i + \epsilon_i \quad (12)$$

$$Sentence_i = \alpha + \beta_1 AppearanceCondition_i + \beta_2 Attractive_i + \beta_3 Arson_i + \beta_4 Eyewitness_i + \epsilon_i \quad (13)$$

$$Sentence_i = \alpha + \beta_1 AppearanceCondition_i + \beta_2 Unattractive_i + \beta_3 Arson_i + \beta_4 Eyewitness_i + \epsilon_i \quad (14)$$

$Sentence_i$ is equal to the assigned sentence length. $AppearanceCondition_i$ is a vector of dummy variables equal to one for five of the six appearance conditions; formal attire with no shackles is the omitted category. $Attractive_i$ and $Unattractive_i$ are equal to the percent of respondents in the second study that find the defendant attractive and unattractive, respectively.⁸¹ $Arson_i$ is a dummy variable equal to one when the crime charged is arson, and $Eyewitness_i$ is a dummy variable equal to one when the case scenario involves eyewitness testimony rather than centering on a motive. Standard errors are clustered at the individual level. The Equation 11 and 12 results are presented in Table 21.

Table 21. Regression Analyses of Defendant Dress, Shackling, and Perceived Attractiveness on Sentencing

	Equation 11	Equation 12
Casual	-0.04	-0.08*
Institutional	0.06	-0.03
Shackles	0.02	-0.001
Attractive	0.001	—
Unattractive	—	0.004
Arson	1.53***	1.53***
Eyewitness	-0.35***	-0.37***
Observations	16,140	16,140
R-squared	0.216	0.216

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

⁸¹ Controlling for attractiveness using the average scores assigned to pictures by a separate sample of respondents in this manner is consistent with other beauty literature (Biddle & Hamermesh 1998). This method assumes common standards of beauty in the population. Previous papers that have examined this assumption have found it to hold (see Biddle & Hamermesh 1998; Hamermesh 2006; Hatfield & Sprecher 1986).

Notes: Standard ordinary least squares regressions are used. The outcome variable is the sentence assigned, between one and five years, inclusive, for vandalism and between one and ten years, inclusive, for arson. Casual and Institutional are dummy variables equal to one if the defendant is in that attire. Shackles is a dummy variable equal to one if the defendant is shackled. Attractive and Unattractive are equal to the percent of respondents in the second study that find the defendant somewhat or very attractive and somewhat or very unattractive on five point Likert scales, respectively. Arson is a dummy variable equal to one if the crime charged is arson. Eyewitness is a dummy variable equal to one if the evidence revolves around eyewitness testimony rather than a motive. Standard errors are clustered at the individual level.

The results in Table 21 illustrate that neither attractiveness nor unattractiveness is significantly correlated with assigned sentence length. When attractiveness is controlled for, defendants in casual attire still receive significantly shorter sentences than defendants in institutional attire, and when unattractiveness is controlled for, defendants in casual attire receive significantly shorter sentences than defendants in formal attire (by 37 and 29 days, respectively).

The Equation 13 and Equation 14 results, which allow the effects of defendant dress and shackles to interact, tell a similar story. When attractiveness is controlled for, defendants in casual attire without shackles receive significantly shorter sentences than defendants in casual attire with shackles and defendants in institutional attire (with or without shackles), defendants in institutional attire without shackles receive significantly longer sentences than defendants in either formal attire with shackles or casual attire with shackles, and defendants in institutional attire with shackles receive significantly longer sentences than defendants in formal attire with shackles. When unattractiveness is controlled for, defendants in casual attire without shackles receive significantly shorter sentences than defendants in formal attire (with or without shackles) and defendants in institutional attire (with or without shackles).

These results suggest that the sentencing effects of defendant dress are not the result of changes in respondents' assessments of attractiveness. Again, this study is not able to tease out the exact mechanism, but assuming decisionmakers are sentencing defendants based on the goals of incapacitation, deterrence, retribution, and rehabilitation, the results suggest that clothing

serves as a signal of culpability or susceptibility to rehabilitation. Discerning the precise source of the bias remains an important area for future research.

VIII. Discussion and Conclusion

The conviction results presented in this Chapter illustrate that certain subgroups of mock jurors are more likely to convict defendants in institutional attire than defendants in civilian attire, and other subgroups of mock jurors are less likely to convict defendants in casual civilian attire than defendants in formal civilian attire. Just as the Court feared in *Estelle*,⁸² the bias against defendants in institutional attire appears to be the result of differences in the presumption of innocence based on either faith in the justice system, a belief that defendants in formal attire have a lower propensity to commit crime, or both, depending on the subgroup.

The effect of casual attire likely stems from the fact that mock jurors interpret the burden of proof to be significantly higher in cases where the defendant is in casual attire rather than formal attire. However, these differences in the numerical interpretation of the burden of proof are not accounted for by reported concerns over the cost of wrongful convictions, the cost of wrongful acquittals, or enforcing the burden of proof threshold. One possible explanation is that these mock jurors feel sympathy for defendants in casual attire because of their financial situation, and they are either unaware of the effect that this sympathy has on their burden of proof interpretation, or they do not report the effect in terms of their concern over the costs of a wrongful conviction.

The casual attire results stand in contrast to the results of Conley, Turnier, and Rose (2000) who did not find a significant effect of casual attire as compared to formal attire on

⁸² 425 U.S. 501 (1976).

verdict decisions when examining the crime of vandalism. The difference in results most likely stems from the fact that Conley, Turnier, and Rose (2000) did not examine the effects of defendant dress on subgroups of mock jurors. The effect of casual clothing on guilt determinations only appears for certain subgroups of mock jurors. It is particularly important to study subgroup effects in this context because important case decisions are typically made by either a single judge or a jury that may contain a large over or underrepresentation of various subgroups depending on the jury selection process and the demographics in the surrounding area.

Additionally, the sentencing results illustrate that mock jurors who convicted the defendant sentence defendants in casual civilian attire the least severely and defendants in institutional attire the most severely, particularly when the defendant is white. The effect of defendant dress alone on sentencing is large: Defendants in casual attire receive about 45 fewer days in prison than defendants in institutional attire. The effect gets even larger when shackles and dress are allowed to interact: Defendants in casual attire without shackles receive about 60 fewer days in prison than defendants in institutional attire (with or without shackles). This significant difference is not the result of differences in recidivism predictions. White defendants in casual attire receive significantly lighter sentences than white defendants in institutional attire from mock jurors who convicted—despite the fact that they are perceived as equally likely to recidivate, regardless of dress. The persistence of the difference in sentencing, after controlling for recidivism predictions, suggests that decisionmakers may associate defendant appearance (and white defendant appearance in particular) with opinions about the defendant's capacity for rehabilitation or the defendant's culpability. Future research should aim to pinpoint the precise mechanism of this effect.

This study focuses on the context of nonviolent property crimes. These crimes account for about one third of all arrests, making them particularly important to study (Federal Bureau of Investigation 2017). However, it is possible that certain features of defendant appearance (shackles in particular) have a negative effect on case outcomes only in other kinds of cases, such as violent crime cases. The only other study to examine the effects of institutional attire on case outcomes may be illustrative on this point, as it focused on the context of a criminal trial for first degree murder (Fontaine & Kiger 1978). The findings suggest that defendant appearance may have a different effect in violent crime cases—defendants in institutional clothing *or* accompanied by an armed guard were worse off than those in personal clothing without an armed guard, but defendants in institutional clothing *also* accompanied by an armed guard were not worse off, apparently because decisionmakers felt sympathetic towards these defendants. It is also possible that shackling influences decisions about what type of sentence to impose but not decisions about the length of an incarceration sentence. Chapter Three will shed some light on this issue by examining the effect of shackling on decisions about what type of sentence to impose in juvenile cases.

As in Chapter One, one potential limitation of this study may be external validity. The mock jurors did not witness a real trial with all of its nuances or a real in-person defendant, nor did they deliberate with others. The existing evidence on the impact of jury deliberation on biases is mixed. Bernard (1979) found white jurors exhibiting bias towards black defendants to be uncompromising during deliberations, and Kramer, Kerr, and Carroll (1990) found that the biasing effect of emotional publicity increased following jury deliberations. Conversely, Sommers (2006) found that deliberations reduce bias against black defendants when the jury is racially mixed, and others (Kerwin & Shaffer 1994, London & Nunez 2000, Ruva et al. 2007)

have found that the biasing effect of inadmissible information is reduced following deliberations. The best evidence that pre-deliberation decisions made by mock jurors are representative of post-deliberation jury decisions comes from a series of mock jury studies by Davis et al. (1975) that compared actual pre-deliberation verdict decisions in criminal cases with the jury verdicts reached after deliberations and found that the best way to transform pre-deliberation verdicts to post-deliberation verdicts is to predict that any verdict favored by two-thirds of the jurors pre-deliberation would be returned by a jury post-deliberation and to predict a hung jury if no such verdict exists.

Additionally, mock jurors may behave differently than real world judges. The sample of mock jurors with the education and income levels of typical judges in this study is small, making it hard to draw clear conclusions about whether bias among them differs significantly from bias among mock jurors as a whole. Moreover, there are ways in which real world judges differ even from mock jurors with similar demographic characteristics. For example, judges are repeat players in the criminal justice system. As such, they can compare defendants and facts across cases. Nevertheless, there is some evidence that judges generally suffer from similar biases as jurors (Guthrie, Rachlinski & Wistrich 2007; Rachlinski & Johnson 2009; Rachlinski, Wistrich, & Guthrie 2015; Wistrich, Rachlinski, & Guthrie 2015).

Even acknowledging these potential limitations, this Chapter provides the first empirical evidence on the effect of shackling and the heterogeneous effects of defendant dress. It is also the first study to explore possible explanations for the effects of defendant dress. Most importantly, this Chapter provides the first empirical support for the Supreme Court's holding in *Estelle*—that defendants should not be forced to appear before a jury in identifiable prison garb. The results illustrate that such attire undermines the presumption of innocence afforded

defendants by certain subgroups of mock jurors. Specifically, the bias against defendants in institutional attire in terms of conviction rates appears primarily among male mock jurors, Republican mock jurors, religious mock jurors, and mock jurors with no criminal history, and the effects are additive, such that mock jurors with all of these characteristics are about 16 percentage points more likely to convict a defendant in institutional attire than a defendant in formal attire.

The fact that this bias is only significant for male mock jurors, Republican mock jurors, religious mock jurors, and mock jurors with no criminal history indicates that the bias is strongest among these subgroups. It does not mean, however, that every individual who falls into one of these subgroups will be biased or that every individual who does not fall into one of these subgroups will not be biased. Rather than providing a blueprint for isolating potential jurors who will be biased in a particular way, subdividing the sample illustrates that bias is present among substantial groups in a way that could influence case outcomes, even though it is not significant for the population as a whole.

While it may be tempting to conclude that male mock jurors, Republican mock jurors, religious mock jurors, and jurors with no criminal history should be struck from juries to prevent the influence of this bias, such behavior would be over-inclusive, under-inclusive, and potentially unconstitutional. Potential jurors can be excused either for cause by a judge or with a preemptory strike by the prosecutor or defendant. Membership in a group, even one that openly advocates a particular position, is typically not enough to justify exclusion for cause,⁸³ and the use of preemptory strikes to intentionally discriminate against members of a particular race or

⁸³ See, e.g., *United States v. Salamone*, 800 F.2d 1216 (1986) (finding that membership in the National Rifle Association was not enough to justify exclusion for cause in a case involving charges under gun control statutes).

sex violates the Equal Protection Clause.⁸⁴ Lower courts are divided about whether the Equal Protection Clause also prevents discriminatory use of preemptory strikes against members of other protected classes, such as religion.⁸⁵

Even if some of these subgroups could constitutionally be targeted for the use of preemptory strikes, it would be more accurate for the parties to weed out bias by asking potential jurors questions related to the source of the bias. For example, a lawyer could ask potential jurors whether there is anything about the defendant's attire that would make it difficult for them to presume that the defendant is innocent before hearing any evidence. An affirmative response would not only provide a solid basis for a preemptory strike but would likely justify removal for cause. Alternatively, the lawyer could ask jurors whether they have faith in the justice system and/or whether they believe crime is concentrated among the poor. The answers to these questions are associated with appearance bias and would provide a less discriminatory ground for preemptory strikes.

The conviction results also illustrate that certain subgroups of mock jurors exhibit a bias in favor of defendants in casual attire. This effect appears primarily among mock jurors with no college education and mock jurors who are unemployed (i.e., mock jurors without the common characteristics of seated jurors). These mock jurors are about 4 percentage points more likely to convict defendants in formal attire than defendants in casual attire. In light of the fact that no subgroup of mock jurors exhibits a counterbalancing bias against defendants in casual attire, these results suggest that, at least in nonviolent criminal cases, the current emphasis on dressing up for trial may be a bit overblown. To the extent that a defense team is deciding whether to

⁸⁴ *J.E.B. v. Alabama ex rel. T.B.*, 511 U.S. 127 (1994) (sex); *Batson v. Kentucky*, 476 U.S. 79 (1986) (race).

⁸⁵ *United States v. Stafford*, 136 F.3d 1109, 1114 (7th Cir. 1998) (“The constitutional status of preemptory challenges based on religion is unsettled”); see also *Davis v. Minnesota*, 511 U.S. 1115 (1994) (denying certiorari in a case raising the issue of the constitutionality of preemptory strikes based on religion).

spend resources on formal attire for trial or something like extra investigation, the latter may have a greater potential impact on case outcomes.

Finally, the results illustrate that white defendants in institutional attire receive the harshest sentences and white defendants in casual attire receive the most lenient sentences from mock jurors who convicted. While considering defendant dress in the context of a sentencing hearing is not necessarily unconstitutional, it may not be ideal from a policy standpoint. Prison attire may signal dangerousness to the community, but it may also signal poverty. Similarly, casual civilian attire in the courtroom may signal a lack of respect for the court process, but it may also be the result of financial inability to secure more formal civilian attire. Policymakers may therefore prefer that judges focus on other, more relevant factors and avoid using defendant dress as a proxy for dangerousness or respect. At least with respect to institutional attire, this objective could be achieved by extending the rights afforded to defendants in *Estelle* to sentencing proceedings. Such a policy would prevent the government from forcing criminal defendants to appear at sentencing proceedings in identifiable prison garb. As in the context of proceedings before a jury, defendants being detained prior to their sentencing would need to be given an opportunity to change into civilian attire before entering the courtroom.

CHAPTER THREE: NO SYMPATHY FOR THE SHACKLED: BANS ON INDISCRIMINATE JUVENILE SHACKLING LEAD TO MORE FAVORABLE CASE OUTCOMES FOR THE ACCUSED

I. Introduction

There is no national consensus regarding juvenile defendants' right to be unshackled in the courtroom. In *Deck v. Missouri*, the Supreme Court extended the right not to be shackled in front of a jury absent a compelling state justification to adult criminal defendants,⁸⁶ but that holding does not automatically apply in the juvenile context. Since the juvenile justice system in the United States has historically been viewed as civil rather than criminal, the same Due Process rights are not always afforded to juvenile defendants.⁸⁷ For the same reason, even if the ruling did apply in juvenile court, juvenile defendants rarely have the right to a jury trial.⁸⁸ As a result, a number of bans on indiscriminate juvenile shackling have been implemented at the state level. Importantly, these bans are not complete bans on shackling, but they limit the circumstances under which juveniles can be shackled. The prohibitions on indiscriminate shackling that exist are based on state statutes, state court rules, state administrative orders, and state court decisions. Table 1 contains a list of the states that have adopted a ban, the source of that ban, and the year the ban was adopted. Today, thirty-two states and the District of Columbia limit juvenile shackling in some way.

⁸⁶ 544 U.S. 622 (2005).

⁸⁷ *In Re Gault*, 387 U.S. 1, 17 (1967) (requiring only that, in proceedings that determine delinquency with a potential consequence of incarceration, Due Process rights that are essential to fundamental fairness be afforded to juveniles).

⁸⁸ *McKeiver v. Pennsylvania*, 403 U.S. 528 (1971) (holding that the right to a jury trial is not constitutionally required in at the adjudication stage of state juvenile cases). The states that do afford juveniles a right to a jury trial are Alaska (ALASKA DELINQUENCY R. 21(a)), Kansas (KAN. STAT. ANN. § 38-2357), Massachusetts (MASS. GEN. LAWS ANN. Ch. 119, § 55A), Michigan (MICH. COMP. LAWS § 712A.17(2)), New Mexico (N.M. STAT. ANN. § 32A-2-16(A)), Oklahoma (OKLA. STAT. ANN. tit. 10, § 7303-4.1), Texas (TEX. FAM. CODE ANN. § 54.03(b)(6)), West Virginia (W. VA. CODE ANN. § 49-5-6(a)), and Wyoming (WYO. STAT. ANN. § 14-6-223(c)).

Table 1. State-Wide Bans on Indiscriminate Juvenile Shackling

State	Source of Ban	Year Adopted
Alaska	ALASKA DELINQ. CT. R. 21.5	2015
Arizona	17B A.R.S. JUV. CT. RULES OF PROC. 12(E)	2017
California	<i>Tiffany A. v. Super. Ct.</i> , 150 Cal. App. 4 th Supp. 1334 (2007)	2007
Connecticut	CONN. GEN. STAT. ANN. § 46b-122a	2015
Delaware	DEL. CODE ANN. Tit. 10, § 1007B	2017
Florida	FLA. R. JUV. P. 8.100(b)	2010
Idaho	<i>State v. Doe</i> , 333 P.3d 858 (Idaho Ct. App. 2014)	2014
Illinois	ILL. SUP. CT. R. 943	2017
Indiana	IND. CODE § 31-30.5-2-1	2015
Kentucky	KY. JUV. R. PRAC. & PROC. 23	2016
Maine	2015 ME Rules 20, Maine Rules of Criminal Procedure, Amended Rule 43A	2015
Maryland	<i>In Re D.M.</i> 228 Md. App. 451 (Md. Ct. Spec. App. 2016); Maryland Judiciary Resolution Regarding Shackling of Children in Juvenile Court	2015
Massachusetts	Trial Ct. of the Commonwealth, CT. OFFICER PL'Y AND PROCEDURES MANUAL, Ch. 4 § VI	2010
Nebraska	NEB. REV. STAT. ANN. § 43-251.03	2015
Nevada	NEV. REV. STAT. ANN. § 62D.415	2015
New Hampshire	N.H. REV. STAT. § 126-U:13	2010
New Jersey	N.J. R. Ch. Div. Fam. Pt. 5:19-4	2017
New Mexico	N.M. CHILD CT. R. 10-223A	2012
New York	N.Y. COMP. CODES R. & REGS. tit. 9, § 168.3(a)	2013
North Carolina	N.C. GEN. STAT. § 7B-2402.1	2007
North Dakota	N.D. R. JUV. P. 20	2017
Ohio	OHIO SUP. R. § 5.01	2016
Oregon	S.B. 846, 2017 Sess. (Or. 2018)	2018
Pennsylvania	237 PA. CODE §139; 42 PA. CONS. STAT. § 6336	2011
South Carolina	S.C. CODE ANN. § 63-19-1435	2014
Tennessee	TENN. R. JUV. PROC. 204	2016
Utah	UTAH CODE ANN. § 78A-6-122	2015
Vermont	33 VT. STAT. ANN. Tit. 33, § 5123	2013
Washington	WASH. JUV. CT. R. 1.6	2014
Washington, D.C.	D.C. Super. Ct. Admin. Order 15-07	2015

Notes: Where there are two sources for a state-wide ban, only the year for the first is listed.

The bans on indiscriminate juvenile shackling that exist are far from symbolic. There are numerous anecdotal accounts documenting the actual practice of indiscriminate juvenile

shackling in a number of states throughout the country.⁸⁹ Proponents of these bans argue that indiscriminate juvenile shackling, like adult shackling, undermines defendants' presumption of innocence and interferes with defendants' right to effective assistance of counsel (Nabha 2008). They also argue that shackling is particularly inappropriate in the juvenile context because it undermines the dignity of the juvenile justice system (which is supposed to center around rehabilitation) and causes psychological harm to juvenile defendants whose brains are still developing (Nabha 2008). The first three of these concerns are similar to the reasoning that the Supreme Court relied on in *Deck* when it held that indiscriminate shackling of adult defendants in front of a jury violates Due Process.⁹⁰ On the other hand, those who oppose bans on indiscriminate shackling argue that they undermine the security of the courtroom (Gallagher & Lore 2008).

This paper will provide the first empirical evidence on the impact of state-wide bans on indiscriminate juvenile shackling. Using difference-in-differences and difference-in-difference-in-differences regression analyses and juvenile criminal case data from North Carolina and Tennessee, obtained from the National Juvenile Court Data Archive, it will illustrate that North Carolina's legislative ban on indiscriminate juvenile shackling resulted in more favorable case outcomes for juvenile defendants, including a 3–4 percentage point decrease in the probability of detention and an equally large increase in the probability of treatment.

In contrast to the experimental results from Chapter Two—finding a very limited effect of shackling in the context of an adult criminal trial for a nonviolent property crime—these

⁸⁹ For example, there are accounts of indiscriminate juvenile shackling in Massachusetts before its ban (McLaurin 2012), in Florida, North Carolina, and North Dakota prior to their bans (Moore 2007), in New York before its ban (Nabha 2008), and in Iowa before its ban (Russell 2017). Additionally, according to a public defender in Colorado (a state with no ban on indiscriminate shackling), juveniles in parts of Colorado were being indiscriminately shackled as recently as 2018.

⁹⁰ 544 U.S. 630–32.

observational results suggest that shackling may in fact influence the outcomes of juvenile criminal cases. The difference in results could indicate that shackling has a unique effect in the juvenile context or could stem from differences in the scopes of the two chapters, which I discuss in more detail below. Whatever the cause, the results in this Chapter illustrate that there are very real consequences of indiscriminate juvenile shackling, providing empirical justification for a nationwide ban on the practice.

II. Literature Review

Apart from Chapter Two of this dissertation, there is no existing empirical literature on the effects of shackling on case outcomes—in the adult or juvenile context. The experimental results from Chapter Two illustrate that in the context of a nonviolent property crime in adult court, shackling generally does not affect the trial verdict or the length of the assigned sentence. Shackling only has an independent significant effect on case outcomes in one instance—mock jurors who identify as Democrat are significantly *less* likely to convict defendants in shackles, suggesting that shackling may actually engender a feeling of sympathy among certain decisionmakers in some contexts. Shackling does have a significant effect on case outcomes when considered in combination with defendant attire, however. Defendants in casual clothing are typically better off than other defendants but usually only when they are not shackled.

There are a number of reasons to suspect that the results in this Chapter may differ from the results in Chapter Two. First, this Chapter examines the effect of shackling across all case types rather than focusing specifically on nonviolent property crimes. The Chapter Two results along with the results from a study by Fontaine and Kiger (1978) suggest that defendant appearance may have a different effect on case outcomes in violent crime cases and property crime cases. Recall that in Fontaine and Kiger (1978), defendants in a simulated murder trial

wearing institutional clothing *or* accompanied by an armed guard were worse off than those in personal clothing without an armed guard, but defendants in institutional clothing *also* accompanied by an armed guard were not worse off—apparently because decisionmakers felt sympathetic towards these defendants. The last Section of this Chapter will specifically address the issue of heterogeneity by crime type.

This Chapter also examines the effect of shackling on a variety of case outcomes rather than on trial verdicts and prison sentence length alone. At an adult criminal trial, decisionmakers are presented with all the evidence available from both sides and, in the case of a jury trial, they are given a number of jury instructions to guide their decisionmaking. Case outcomes such as dismissals occur in a much less structured environment, where shackling may have a different effect. This Chapter also examines categorical sentencing decisions rather than prison sentence length decisions. Shackling may have an effect on whether or not a judge feels it is appropriate to assign detention or some other type of sentence but not on the length of a prison sentence conditional on the defendant being incarcerated.

Additionally, Chapter Two only manipulated the presence or absence of shackles in a still picture of the defendant. All other case-related information was held constant. This Chapter examines a state-wide ban on indiscriminate juvenile shackling, which may influence case outcomes not only by modifying the appearance of juvenile defendants but also by changing the content of the adjudication itself, for example, by allowing juvenile defendants to more effectively communicate with their attorneys or making juvenile defendants feel more comfortable so they can more effectively advocate on their own behalf.

Finally, the most obvious difference between this Chapter and Chapter Two is that this Chapter focuses on the juvenile context. It is possible that the sight of a child in shackles has a

different effect on decisionmakers than the sight of an adult in shackles. A shackled child could appear more innocent and sympathetic than a shackled adult, or the sight of a shackled juvenile could be more jarring than the sight of a shackled adult, making the child appear more threatening and dangerous.

The empirical literature most closely related to this project is the appearance literature discussed in detail in Chapter Two. Outside the context of the courtroom, this literature has illustrated that plain looking people earn less than the good-looking and that this effect is largely independent of occupation. These findings appear to be the result of taste-based employer discrimination—i.e., discrimination stemming from decisions made by persons deriving disutility from interacting with less attractive individuals (Becker 1957; Hamermesh & Biddle 1994). Some of the gap in wages is also explained by the fact that attractive workers are more confident, and confidence increases wages; for a given level of confidence, attractive workers are (wrongly) considered to be more able by employers; and controlling for confidence, attractive workers have communication and social skills that raise their wages when they interact with employers (Mobius & Rosenblat 2005). Beauty has also been shown to impact happiness (Gupta, Etcoff, & Jaeger 2015; Hamermesh & Abrevaya 2013), political elections (Berggren, Jordahl, & Poutvaara 2017; Jones & Price 2017; Leigh & Susilo 2009), charitable giving (Cryder, Botti, & Simonyan 2017; Price 2008), instructor ratings (Hamermesh & Parker 2005), profits (Rule & Ambady 2008), academic performance (Hernandez-Julian & Peters 2017), and propensity to commit crime (Mocan & Tekin 2010).

In the criminal context, most experimental studies have found that defendant beauty impacts guilt (Abwender & Hough 2001; Desantis & Kayson 1997; Efran 1974; Leventhal & Kreate 1977; MacCoun 1990; Patry 2008) and sentencing decisions (Abwender & Hough 2001;

Efran 1974; Friend & Vinson 1974; Leventhal & Kreate 1977; Piehl 1977; Sigall & Ostrove 1975; Smith & Hed 1997). Typically, more attractive defendants are better off, but a meta-analysis by Mazzella and Feingold (1994) found that the effect of beauty on both guilt determinations and sentencing is small, and the sentencing effects of beauty differ based on the crime charged. Attractive defendants receive less punishment for robbery, rape, and cheating; more punishment for negligent homicide; and the same amount of punishment as unattractive defendants for swindling. Chapter Two also illustrates that defendant attire can affect case outcomes at trials for nonviolent property crimes—defendants in institutional attire are convicted the most and receive the harshest sentences, and defendants in casual clothing are convicted the least and receive the lightest prison sentences.

A couple of studies have examined the effect of other visual cues on case outcomes and juror stereotypes. In a recent experimental study in Australia, Rossner et al. (2017) found that placement of the defendant in an open dock, a glass dock, or at the bar table next to their attorney can affect trial outcomes: Jurors are more likely to convict defendants when they are in a traditional or secure dock. Lassiter (2002) uses an experiment to show that videotaped confessions where the camera is focused on the suspect as compared to videotaped confessions where the camera is focused on both the suspect and interrogator are judged to be more voluntary and lead to more guilty verdicts.

Finally, there are a number of anecdotal accounts about how the presence of shackles affects the attitudes and behavior of juvenile defendants in the courtroom. A juvenile judge in Massachusetts noted, “[Ending indiscriminate shackling] has improved the atmosphere and the culture of the courtroom. When a child can turn and actually say ‘hello,’ and you can see somebody smile back, that changes things for the child and the family member. It also makes it

easier for the management of the courtroom (National Juvenile Defender Center).” Another juvenile judge in New Hampshire noted, “Automatically restraining a juvenile in the courtroom deprives that young person of the opportunity to show the court they are capable of self-control. To the extent that restraints result in an unnatural posture or gait, or a submissive demeanor, a juvenile may feel wholly defeated as he or she enters the courtroom (National Juvenile Defender Center).” In an affidavit addressing indiscriminate juvenile shackling and its potential psychological costs, child psychologist Dr. Marty Beyer wrote, “In the midst of their identity and moral development, demeaning treatment by adults may solidify adolescents’ alienation, send mixed messages about the justice system, and confirm their belief that they are all bad, all of which undermine the rehabilitative goal of court intervention” (Beyer 2015).

A number of other experts have also commented on the tendency of shackling to make juvenile defendants feel anxious, shameful, and humiliated (Bidwell 2015; Ford 2014; Griffin 2014; Kraus 2015; Rosenblitt 2015; Wurm 2015). Differences in behavior could lead to differences in case treatment. For example, prosecutors may be more willing to informally resolve the case of a child who appears engaged with their family and shows self-restraint in the courtroom. Additionally, differences in emotion can affect case outcomes. Defensiveness and anxiety can be detected by others, such as judges, jurors, prosecutors, and defense attorneys, and these decisionmakers may have difficulty distinguishing anxiety and shame related to shackling from anxiety and shame related to guilt (Harrigan et al. 1996). This study will provide the first empirical evidence on the effect of indiscriminate juvenile shackling on case outcomes by examining the effect of a legislative ban on indiscriminate juvenile shackling in North Carolina. The next Section provides some background information on the juvenile justice system and its unique goals and procedures.

III. Background on Juvenile Justice in the United States

The juvenile justice system in the United States revolves around rehabilitation. It was founded at the end of the nineteenth century based on the common sentiments that children who commit crimes are less blameworthy than adults, have a larger capacity for rehabilitation, and are less deserving of harsh punishments commonly imposed in adult court (Feld 2014; Frank 2017). Over time, abuses of the informality of the juvenile justice system resulted in the implementation of procedural protections resembling those of the adult system, but rehabilitation has remained the foremost ideal.⁹¹

Procedurally, there are a few ways that the juvenile justice systems in each state function differently from the adult criminal justice system. After citation or arrest, juvenile defendants go through a process called intake. Intake is an informal procedure that provides an opportunity for cases to be reviewed and either dismissed or handled informally before formal charges are filed (Birkhead 2013; Frank 2017). The intake officer is typically a juvenile probation officer, but some jurisdictions authorize prosecutors to serve as intake officers (Birkhead 2013; Frank 2017). The intake officer will typically meet with the juvenile and ask him or her a series of questions about what happened and about the juvenile's background and criminal history (Birkhead 2013; Frank 2017). The intake officer may have access to additional relevant information, such as a complaint statement or the child's school records (Birkhead 2013; Frank 2017). After the

⁹¹ In the 1960s, the Supreme Court began formalizing the juvenile justice system, because it recognized the need for procedures to protect the rights of juvenile defendants. *See In re Winship*, 397 U.S. 358, 368 (1970) (holding that proof beyond a reasonable doubt is a requirement during the adjudicatory stage of a delinquency proceeding); *In re Gault*, 387 U.S. at 31–59 (holding that, in proceedings that determine delinquency with a potential consequence of incarceration, the right to notice, the right to counsel, the right to confrontation, and the right against self-incrimination will be afforded to juveniles). In 1980, Congress also got involved, aiming to remedy flagrant abuses of power during the detention stage by passing the Criminal Rights of Institutionalized Persons Act, which gave the Attorney General authority to remedy “egregious or flagrant conditions” that deprive inmates of their “rights, privileges, or immunities secured or protected by the Constitution or laws of the United States.” 42 U.S.C. § 1997(a) (2018).

meeting, the intake officer will decide whether to (1) dismiss the case; (2) deal with it informally—usually by having the child consent to a period of probation, having the child consent to some other resolution such as paying restitution, or by referring the child to a diversion program involving an education or treatment program; or (3) recommend the filing of formal charges, a process known as petitioning (Birkhead 2013; Frank 2017).

Juveniles can be charged with delinquency offenses or status offenses. Delinquency offenses involve actions that would also be criminal if the juvenile was an adult at the time. Status offenses, on the other hand, involve actions that are only criminal for juveniles. Underage possession of tobacco and running away are examples of status offenses, and trespassing and murder are examples of delinquency offenses.

Once formal charges have been filed in a case, the juvenile court judge must determine whether they have jurisdiction to hear the case in juvenile court, and, in most states, juvenile court judges may waive juvenile court jurisdiction in certain cases and elect to have those cases transferred to adult court (Freitas 1995; Hockenberry & Puzanchera 2018). Legislative limits on the jurisdiction of juvenile courts are typically based on the age of the juvenile and the seriousness of the offense (Freitas 1995). Judicial decisions to waive jurisdiction are typically based on a determination about whether the juvenile is amenable to treatment in the juvenile system (Freitas 1995). A judge might find, for example, that a juvenile who has been adjudicated several times in the past is not amenable to treatment or that a juvenile charged with a serious offense is not able to be properly rehabilitated in the time period that the juvenile court would have jurisdiction over him or her.

If the juvenile court determines that it has jurisdiction to hear the case and decides not to transfer the case to adult court, then the case proceeds in a manner similar to cases in the adult

criminal justice system. The juvenile may be detained prior to case resolution if there is reason to believe that he or she is a threat to the community, will be at risk if returned to the community, or may fail to appear at an upcoming hearing if released (Sickmund 2005). Juveniles may engage in plea bargaining, and the option of informal case resolution—avoiding a conviction through voluntary probation, restitution, or participation in treatment—remains available (Sickmund 2005).

If no agreement can be reached between the parties, an adjudicatory hearing is held to establish responsibility for the alleged act (Sickmund 2005). A juvenile defendant being adjudicated delinquent is equivalent to an adult defendant being found guilty. However, as mentioned earlier, juveniles typically do not have the right to a trial by jury, so adjudicatory hearings are almost always presided over by juvenile judges.⁹² Finally, once a juvenile is adjudicated delinquent, the court assigns a disposition (or sentence) (Sickmund 2005).

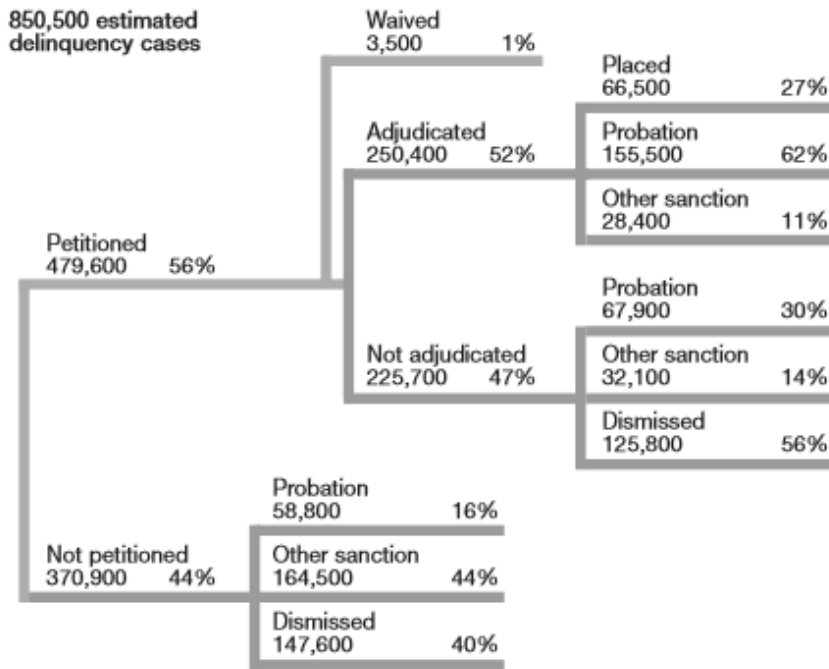
Dispositions can involve educational or treatment programs designed to assist with rehabilitation; probation; secure detention in a juvenile detention facility; or some other punishment, such as a fine, an order to pay restitution, or an order to complete a certain amount of community service.

Figure 1 depicts juvenile case processing as a flow chart and provides summary statistics regarding juvenile case resolutions across the United States in 2016 (Hockenberry & Puzzanchera 2018). Fifty-six percent of cases in 2016 resulted in the filing of formal charges. Of those cases that were petitioned, only 1 percent were judicially transferred to adult criminal

⁹² The states that do afford juveniles a right to a jury trial are Alaska (ALASKA DELINQUENCY R. 21(a)), Kansas (KAN. STAT. ANN. § 38-2357), Massachusetts (MASS. GEN. LAWS ANN. Ch. 119, § 55A), Michigan (MICH. COMP. LAWS § 712A.17(2)), New Mexico (N.M. STAT. ANN. § 32A-2-16(A)), Oklahoma (OKLA. STAT. ANN. tit. 10, § 7303-4.1), Texas (TEX. FAM. CODE ANN. § 54.03(b)(6)), West Virginia (W. VA. CODE ANN. § 49-5-6(a)), and Wyoming (WYO. STAT. ANN. § 14-6-223(c)).

court, 52 percent resulted in a delinquency adjudication, and 47 percent were either dismissed or handled informally.

Figure 1. Juvenile Case Processing in the United States in 2016



Notes: This image is borrowed from the National Center for Juvenile Justice's 2016 Juvenile Court Statistics Report. The report uses data obtained from the National Juvenile Court Data Archive. Cases are categorized by their most severe or restrictive sanction, and detail may not add to totals because of rounding. Annual case processing flow diagrams for 1985 through 2016 and additional methodological information can be found online at ojjdp.gov/ojstatbb/court/faqs.asp.

IV. Methodology

This Chapter will empirically estimate the effect of a ban on indiscriminate juvenile shackling on juvenile case outcomes using state-level juvenile case data obtained from the National Juvenile Court Data Archive. Specifically, the study will compare case outcomes for juveniles who were and were not incarcerated during the pendency of their cases in North Carolina and Tennessee before and after North Carolina adopted a state-wide legislative ban on indiscriminate juvenile shackling in 2007. The next parts will discuss the methodology that will

be used, provide background on the juvenile justice procedures and relevant laws in North Carolina and Tennessee, and describe the data that will be used.

A. Double-Difference and Triple-Difference Regression Analyses

Using traditional ordinary least squares regression analyses to investigate the impact of indiscriminate juvenile shackling on case outcomes would likely produce biased results, because—even were case-level shackling data available—it would be impossible to control for all of the case-specific facts that are correlated with shackling and with case outcomes.

Traditional ordinary least squares regression analyses would also raise reverse causality concerns, as juveniles with higher probabilities of guilt and greater culpability may be more likely to be shackled—even though shackling is often indiscriminate. In general, the only juveniles being shackled are those who are incarcerated during the pendency of their proceedings. To avoid these concerns and identify a causal effect, this Chapter will use a two-state difference-in-differences (hereinafter, double-difference) analysis and a two-state difference-in-difference-in-differences (hereinafter, triple-difference) analysis.

In essence, the two-state double-difference analysis utilizes a state without a ban on indiscriminate juvenile shackling as a control and compares it to a state that has adopted a ban on indiscriminate juvenile shackling before and after the passage of the ban. The comparison over time and between the two states nets out changes in outcomes attributable to shared experiences of the states and differences in outcomes based on time-invariant differences between the states. As long as (1) there was no other discrete change that occurred around the same time as the passage of the ban that would influence juvenile case outcomes in one state and not the other, and (2) the outcomes of interest would have progressed in a parallel manner in the two states absent the passage on the ban (an assumption often referred to as the parallel trends assumption),

the double-difference estimates will represent the causal effect of the ban on juvenile case outcomes, alleviating concerns about omitted variables and reverse causality.

In order to isolate the effects of North Carolina’s ban on indiscriminate juvenile shackling, outcomes from Tennessee (a state without a statutory ban on indiscriminate juvenile shackling during the relevant time period) will be compared to outcomes in North Carolina (a state that adopted a statutory ban on indiscriminate juvenile shackling during the relevant time period) between 2004 to 2012, before and after North Carolina’s law became effective in 2007.⁹³ Equation 1 is the main regression equation of interest.

$$Outcome_{ict} = \alpha + \beta_1 CaseControls_{ict} + \beta_2 County_c + \beta_3 Month_t + \beta_4 County_c * Trend_t + \beta_5 NorthCarolina_s * PostSept2007_t + \epsilon_{ict} \quad (1)$$

$Outcome_{ict}$ is a dummy variable equal to one if the case resulted in the particular outcome being examined. $CaseControls_{ict}$ is a vector of controls for case-level characteristics, discussed in more detail below. $County_c$ and $Month_t$ are county- and month-level fixed effects, and $NorthCarolina_s$ and $PostSept2007_t$ are indicator variables for whether a case was in North Carolina and after September 2007, respectively. County specific linear time trends are also included by interacting dummy variables for each county with the variable $Trend_t$ that equals one in 2004, two in 2005, and so on. These trends control for a variety of other variables that might influence case outcomes (such as the tenure of juvenile court judges), but which trend smoothly over time. β_4 will represent the causal effect of the ban on indiscriminate shackling on case outcomes as long as the two identifying assumptions listed above hold true.

⁹³ For an example of a paper utilizing a similar two-state double-difference model, see Card & Krueger (1994), which calculates the impact of an increase in the minimum wage on employment in the fast food industry by comparing total employment in fast food restaurants in New Jersey and Pennsylvania before and after the minimum wage was increased in New Jersey.

The first identifying assumption—that there was no other discrete change that occurred around the same time as the passage of North Carolina’s ban on indiscriminate juvenile shackling that would influence juvenile case outcomes in either North Carolina or Tennessee—is likely met in this this instance. There are three primary sources of potential changes: state juvenile court rules, state juvenile codes, and state court opinions—none of which reveal any discrete changes in 2007 that seriously undermine the double-difference methodology.⁹⁴

In 2007, there were no juvenile court rule update orders filed in North Carolina. Besides the legislative ban on indiscriminate juvenile shackling, there were five other amendments to North Carolina’s juvenile code: (1) an act to amend existing child welfare laws to comply with federal law and regulations, which, for example, replaces “an opportunity” with “the right” as related to the ability of certain parties to be heard and present evidence on issues relating to juvenile placement;⁹⁵ (2) an act to require the department of juvenile justice and delinquency prevention to release the identification of certain juveniles who escape from custody;⁹⁶ (3) an act to provide that the court may order secure custody of a juvenile when the juvenile is charged with driving while ability impaired or underage drinking and to authorize the legislative research commission to study dispositional alternatives for juveniles who are adjudicated delinquent for these offenses;⁹⁷ (4) an act to expand the reach of North Carolina courts in proceedings to

⁹⁴ The information presented regarding these changes is based on Westlaw search results. For state juvenile court rule updates, I searched for the term “juvenile” in rule update orders from January 1, 2007 – December 31, 2007 in both Tennessee and North Carolina. For state juvenile code updates, I searched for the term “juvenile” in historical enacted legislation from January 1, 2007 – December 31, 2007 in both North Carolina and Tennessee. Finally, for state supreme court opinions, I searched for the term “juvenile” in North Carolina and Tennessee Supreme Court cases issued from January 1, 2007 – December 31, 2007. There are, of course, other discrete changes that could have occurred in North Carolina or Tennessee in 2007. My interviews with individuals involved in the juvenile justice systems in North Carolina and Tennessee did not reveal any such events.

⁹⁵ H.B. 698, 2007 Gen. Assemb., 1st Sess. (N.C. 2007).

⁹⁶ H.B. 1148, 2007 Gen. Assemb., 1st Sess. (N.C. 2007).

⁹⁷ S.B. 999, 2007 Gen. Assemb., 1st Sess. (N.C. 2007).

terminate the parental rights of nonresident parents of resident children;⁹⁸ and (5) an act to remove barriers to adoption for residents of other states seeking to adopt children in North Carolina under the laws pertaining to termination of parental rights and adoption.⁹⁹

Additionally, in 2007, the North Carolina Supreme Court issued two opinions related to juvenile justice. In one opinion, the court held that an aunt who did not have custody of the juvenile and with whom the juvenile had not lived for an extended period of time did not qualify as a “guardian” entitled to be present upon the juvenile’s request during police questioning.¹⁰⁰ In the other opinion, the court held that consensual oral sex between a 14-year-old and a 12-year-old violated North Carolina’s crime against nature statute and that the crime against nature statute was not unconstitutional as applied to those juveniles.¹⁰¹

In 2007, there were also no juvenile court rule update orders filed in Tennessee. However, there were eight amendments to Tennessee’s juvenile code: (1) an act to amend the code relative to adoption and guardianship of minors;¹⁰² (2) an act that allows for the release of certain records related to child sexual abuse in certain instances involving claims under the Criminal Injuries Compensation Act;¹⁰³ (3) an act that makes a parent liable for the offense of contributing to the delinquency of a minor if their child is found delinquent for a second or subsequent time of vandalism and the property is owned by the government;¹⁰⁴ (4) an act that requires notifying the juvenile’s school when school attendance is made a condition of probation;¹⁰⁵ (5) an act to amend the code relative to judicial proceedings affecting juveniles,

⁹⁸ H.B. 866, 2007 Gen. Assemb., 1st Sess. (N.C. 2007).

⁹⁹ H.B. 865, 2007 Gen. Assemb., 1st Sess. (N.C. 2007).

¹⁰⁰ State v. Oglesby, 361 N.C. 550 (2007).

¹⁰¹ In Re R.L.C., 361 N.C. 287 (2007).

¹⁰² H.B. 321, 105th Gen. Assemb., Reg. Sess. (Tenn. 2007).

¹⁰³ S.B. 1568, 105th Gen. Assemb., Reg. Sess. (Tenn. 2007).

¹⁰⁴ H.B. 1167, 105th Gen. Assemb., Reg. Sess. (Tenn. 2007).

¹⁰⁵ H.B. 594, 105th Gen. Assemb., Reg. Sess. (Tenn. 2007).

which adds references to the rules of juvenile procedure to clarify that nothing in the law supersedes those procedures;¹⁰⁶ (6) an act that makes court petitions and orders subject to public disclosure in the case of certain serious offenses and establishes a task force to research and handle issues related to the submission of juvenile fingerprints and the reporting of juvenile court dispositions;¹⁰⁷ (7) an act that allows police officers to film or record juveniles they believe are involved in a crime for use as evidence,¹⁰⁸ and (8) an act that requires state agencies administering funds related to the prevention, treatment, or care of delinquent juveniles to only expend funds on evidence-based programs and contract with providers of services to only engage in evidence-based practices.¹⁰⁹ While this last amendment could influence case outcomes by affecting the availability of treatment programs for juveniles, the requirement was phased in over the course of four years in order to prevent that very issue.

In 2007, the Tennessee Supreme Court issued only one opinion related to juvenile justice. The court held that a directive by a juvenile court judge imposing house arrest on a juvenile was invalid because it was never reduced to writing.¹¹⁰ The fact that none of these legal changes seem likely to influence system-wide case outcomes for juvenile defendants in either state suggests that the double-difference estimates represent the casual effects of North Carolina's ban on indiscriminate juvenile shackling on case outcomes as long as the parallel trends assumption is met. The validity of that assumption is discussed below.

The triple-difference methodology compares outcomes for juveniles incarcerated pretrial—making them susceptible to indiscriminate juvenile shackling practices—and juveniles

¹⁰⁶ H.B. 320, 105th Gen. Assemb., Reg. Sess. (Tenn. 2007).

¹⁰⁷ S.B. 2194, 105th Gen. Assemb., Reg. Sess. (Tenn. 2007).

¹⁰⁸ H.B. 75, 105th Gen. Assemb., Reg. Sess. (Tenn. 2007).

¹⁰⁹ S.B. 1790, 105th Gen. Assemb., Reg. Sess. (Tenn. 2007).

¹¹⁰ *State v. Rodgers*, 235 S.W.3d 92 (Tenn. 2007).

not incarcerated pretrial—who would not have been shackled, regardless of the presence of a ban—in North Carolina and Tennessee before and after the passage of North Carolina’s ban in 2007. The main regression equation of interest is:

$$\begin{aligned}
 Outcome_{ict} = & \alpha + \beta_1 CaseControls_{ict} + \beta_2 County_c + \beta_3 Month_t + \\
 & \beta_4 PreTrialDetention_{ict} + \beta_5 County_c * Month_t + \beta_6 County_c * \\
 & PretrialDetention_{ict} + \beta_7 Month_t * PretrialDetention_{ict} + \\
 & \beta_8 NorthCarolina_s * PostSept2007_t * PretrialDetention_{ict} + \\
 & \epsilon_{ict}
 \end{aligned} \tag{2}$$

Again, $Outcome_{ict}$ is a dummy variable equal to one if the case resulted in the particular outcome being examined. $CaseControls_{ict}$ is a vector of controls for case-level characteristics, discussed in more detail below. $County_c$ and $Month_t$ are county- and month-level fixed effects, and $NorthCarolina_s$ and $PostSept2007_t$ are indicator variables for whether a case was in North Carolina and after September 2007, respectively. $PretrialDetention_{ict}$ is a dummy variable equal to one if the juvenile was held in detention pretrial in Tennessee and if the juvenile was charged with a felony offense in North Carolina. The different treatment for cases in the two states is based on a data limitation. Only the Tennessee data contain information about whether juveniles are detained pretrial. Having a felony charge is used to proxy for pretrial detention in North Carolina because juvenile defendants there are more likely to be detained when they are charged with a serious offense.¹¹¹

This analysis increases the precision of the estimates by reporting the effect on the juvenile defendants most likely to be impacted by the ban, rather than the average effect over the whole population of juvenile defendants. Regardless of a change in the law, juveniles who were

¹¹¹ N.C. GEN. STAT. § 7B-1903(b) (stating that a juvenile may be held in secure custody if he or she is charged with a felony or certain misdemeanors and has demonstrated that he or she is a danger to property or persons, he or she has willfully failed to appear on a pending delinquency charge, there is reasonable cause to believe he or she will not appear in court, he or she is an absconder from a residential facility, or there is reasonable cause to believe that he or she should be detained for his or her own protection).

not detained pre-trial would never have been shackled. The triple-difference analysis also helps ensure that any observed changes are indeed the result of North Carolina's legislative ban on indiscriminate juvenile shackling: As long as (1) there was no discrete change that occurred around the same time as the ban that differentially affected juvenile defendants incarcerated pretrial in either North Carolina or Tennessee, and (2) the parallel trends assumption is met, β_8 will represent the causal effect of North Carolina's ban on indiscriminate juvenile shackling on case outcomes for juvenile defendants who may have been incarcerated pretrial.

B. Tennessee as a Control State for North Carolina

Tennessee was selected as the comparator state for North Carolina because as geographic neighbors, they are likely to have shared cultural characteristics and be impacted similarly by economic and ecological shocks. While North Carolina is more populous than Tennessee (by about 3 million individuals) and slightly more diverse (Tennessee is 79 percent white-only and North Carolina is 71 percent white-only), the two states have very similar education levels (both states have an 86 percent high school graduation rate among those 25 or older), median incomes (about \$47,000 per year in Tennessee and \$48,000 per year in North Carolina), and average commute times (24 minutes in North Carolina and 25 minutes in Tennessee) (U.S. Census Bureau b, c). Additionally, both states are predominately Republican with some predominately Democrat urban areas.¹¹² In Tennessee, only two of the nine congressional districts are represented by a Democrat—the 5th Congressional district, which is home to the most populous city in the state (Nashville), and the 9th Congressional District, which is home to the second most

¹¹² Although North Carolina currently has a Democratic governor, both states' electoral college votes went to President Trump in the 2016 presidential election, both states have a majority Republican state house and senate, and both states have entirely Republican U.S. senators and majority Republican U.S. house representatives.

populous city in the state (Memphis). In North Carolina, only three of the thirteen congressional districts are represented by a Democrat—the 12th Congressional District, which is home to most populous city in the state (Charlotte), the 4th Congressional District, which is home to two-thirds of the second most populous city in the state (Raleigh), and the 1st Congressional District, which is home to the fourth most populous city in the state (Durham).¹¹³

Of course, since this paper examines juvenile case outcomes, it is important to understand the unique aspects of North Carolina and Tennessee’s juvenile justice systems. The biggest difference between the two juvenile justice systems is their jurisdictional boundaries. In North Carolina, juvenile courts have jurisdiction to hear delinquency cases involving acts committed by individuals between the ages of 6 and 15, inclusive, and status offense cases involving acts committed by individuals between the ages of 6 and 17, inclusive.¹¹⁴ North Carolina juvenile courts do not have jurisdiction to hear cases involving juveniles who have been emancipated or juveniles who have previously been convicted of a crime as an adult.¹¹⁵ If a juvenile is 13 years of age or older at the time they commit an offense that would be a Class A felony in adult criminal proceedings, the court must transfer jurisdiction to adult court upon a finding of probable cause.¹¹⁶ If a juvenile is 13 years of age or older at the time they commit an offense that would be a lesser felony offense in adult criminal proceedings, the juvenile court judge has the option to transfer the case to adult court.¹¹⁷

¹¹³ There do not appear to have been any major political changes in North Carolina or Tennessee in 2007 that could bias the results. Neither state had a major election in 2007.

¹¹⁴ See N.C. GEN STAT. § 7B1501, 7B-1600–1601.

¹¹⁵ *Id.*

¹¹⁶ N.C. GEN STAT. § 7B-2200. Offenses that would constitute a Class A felony in adult criminal proceedings include first degree murder, unlawfully and willfully injuring another by the use of a nuclear, biological, or chemical weapon of mass destruction, and willfully and maliciously causing the death of an unborn child.

¹¹⁷ N.C. GEN STAT. § 7B-2203.

In Tennessee, juvenile courts have jurisdiction to hear delinquency and status offense cases involving acts committed by individuals under the age of 18.¹¹⁸ As in North Carolina, Tennessee juvenile courts do not have jurisdiction to hear cases involving juveniles who have previously been convicted of a crime as an adult.¹¹⁹ A juvenile court judge can elect to transfer a case to adult court if: (1) the juvenile was less than 14 at the time of the offense and is charged with first or second degree murder or attempted first or second degree murder, (2) the juvenile was at least 14 years old and younger than 17 years old at the time of the offense and is charged with “first degree murder, second degree murder, rape, aggravated rape, rape of a child, aggravated rape of a child, aggravated robbery, especially aggravated robbery, aggravated burglary, especially aggravated burglary, kidnapping, aggravated kidnapping, especially aggravated kidnapping, commission of an act of terrorism, carjacking, or an attempt to commit any such offenses,” (3) if the juvenile was 16 years or older at the time of the offense and is charged with robbery or attempted robbery, or (4) if the juvenile was 17 years or older at the time of the offense.¹²⁰

Despite these jurisdictional differences, the juvenile justice systems in the two states are similar in many other ways. In both states, the juvenile defense systems are organized at the local rather than state level.¹²¹ Both states have intake procedures and the option for diversion both before and after formal charges have been filed.¹²² Both states allow for secure detention of

¹¹⁸ TENN. CODE ANN. § 37-1-102–103.

¹¹⁹ TENN. CODE ANN. § 37-1-134(c).

¹²⁰ TENN. CODE ANN. § 37-1-134(a).

¹²¹ NATIONAL JUVENILE DEFENSE CENTER, *State Profiles: North Carolina* (July 2018), <http://njdc.info/practice-policy-resources/state-profiles/northcarolina/>; NATIONAL JUVENILE DEFENSE CENTER, *State Profiles: Tennessee* (August 2018), <http://njdc.info/practice-policy-resources/state-profiles/tennessee/>. This could result in more differences in shackling practices by county, because different defenders may be challenging shackling with different frequencies and strategies. Heterogeneity in county shackling practices is discussed more below.

¹²² N.C. GEN STAT. § 7B-1700, 7B-1706; TENN. CODE ANN. § 37-1-110.

juveniles during the pendency of their cases,¹²³ and both states utilize risk assessment instruments to make disposition recommendations, develop probation case plans, and, in the case of North Carolina, inform intake decisions.¹²⁴ In North Carolina, the use of these instruments is mandated by statute,¹²⁵ and in Tennessee, the use of a risk assessment instrument is mandated by Department of Juvenile Justice policy.¹²⁶

Generally, North Carolina and Tennessee appear to be good comparator states. In order to ensure that the main difference in their juvenile justice systems—the age-based definition of a juvenile—is not driving the results discussed below, some of the analyses restrict the sample in both states to juveniles between the ages of 6 and 15 as a robustness check.¹²⁷

C. North Carolina’s Legislative Ban on Indiscriminate Juvenile Shackling

On June 20, 2007, the North Carolina legislature passed a law banning the use of indiscriminate shackling. The law, which became effective on October 1, 2007, states

At any hearing . . . the judge may subject a juvenile to physical restraint in the courtroom only when the judge finds the restraint to be reasonably necessary to maintain order, prevent the juvenile's escape, or provide for the safety of the courtroom. Whenever practical, the judge shall provide the juvenile and the juvenile's attorney an opportunity to be heard to contest the use of restraints before the judge orders the use of restraints. If restraints are ordered, the judge shall make findings of fact in support of the order.¹²⁸

This statutory language is fairly representative of other legislative and administrative bans on indiscriminate juvenile shackling.

¹²³ N.C. GEN STAT. § 7B-2204(b); TENN. CODE ANN. § 37-1-114.

¹²⁴ JUVENILE JUSTICE: GEOGRAPHY, POLICY, PRACTICE & STATISTICS, *North Carolina: Juvenile Justice Services*, <http://www.jjgps.org/juvenile-justice-services/northcarolina/>; JUVENILE JUSTICE: GEOGRAPHY, POLICY, PRACTICE & STATISTICS, *Tennessee: Juvenile Justice Services*, <http://www.jjgps.org/juvenile-justice-services/tennessee/>.

¹²⁵ N.C. GEN STAT. § 7B-2501.

¹²⁶ JUVENILE JUSTICE: GEOGRAPHY, POLICY, PRACTICE & STATISTICS, *Tennessee: Juvenile Justice Services*, <http://www.jjgps.org/juvenile-justice-services/tennessee/>.

¹²⁷ Age at the time of the offense is only available in North Carolina. Therefore, this restriction is based on age at the time of referral, which tends to be very similar to age at the time of the offense.

¹²⁸ N.C. GEN STAT. § 7B-2402.1.

In order for the comparison of North Carolina and Tennessee around the time of passage of this law to be meaningful, the law must have created a real change in the practice of indiscriminate juvenile shackling in North Carolina as compared to Tennessee. In other words, there must have been at least some counties in North Carolina that were indiscriminately shackling juveniles before the law became effective, the law must have been enforced such that indiscriminate juvenile shackling ended in those counties after the passage of the law, and there must have been at least some counties in Tennessee engaging in indiscriminate juvenile shackling between 2004 and 2012. In order to investigate these issues, I conducted informal interviews with lawyers practicing juvenile defense in both states and with juvenile court counselors in North Carolina. While I was not able to come up with a comprehensive list of which counties in which states were shackling juveniles when, I did gain some valuable insight about the source of North Carolina's law, its effect, and the practice of indiscriminate juvenile shackling in Tennessee.

The passage of the ban on indiscriminate juvenile shackling in North Carolina was sparked by a juvenile delinquency case in Greensboro, North Carolina, which is located in Guilford County. A juvenile defendant was being shackled, and Legal Aid of North Carolina got involved out of fear that the shackling was traumatizing the child due to a history of abuse. The result was proposed legislation banning indiscriminate juvenile shackling, which ultimately passed with very little opposition. This legislative history suggests that the law was not symbolic—it was a response to actual shackling practices in the state. However, shackling policies often differ by county, as each county has its own juvenile court and juvenile court judge(s).

Before the law went into effect, there may have been some counties that did not indiscriminately shackle juveniles, and after the law went into effect, counties may have implemented it differently. For example, the juvenile court judge in Cumberland county (which includes the city of Fayetteville) currently holds a hearing in every case to determine whether the wrist and ankle shackles on a juvenile should be removed. In Mecklenburg (which includes the city of Charlotte) and Wake (which includes the city of Raleigh) counties, juveniles are presumptively brought into court without shackles unless there is a motion by the state requesting otherwise, and the juvenile court judge in Durham county (which includes the city of Durham) was publically criticized last year for continuing to engage in indiscriminate juvenile shackling, apparently in contravention of the state-wide ban (Bridges 2018). This qualitative information indicates that the effect of the law was not universal across the state, but there was an effect in some parts of the state. This means that the results presented here will underestimate the true effect of a meaningful ban on indiscriminate juvenile shackling.

The practice of indiscriminate juvenile shackling was also not universal across Tennessee during the time period examined here. For example, in Knox county (which includes the city of Knoxville), juvenile defendants were indiscriminately brought into court in ankle shackles but were never shackled at the wrists. In Hamilton county (which includes the city of Chattanooga), juvenile defendants were indiscriminately shackled at both the ankles and wrists, and in Shelby county (which includes the city of Memphis), juvenile defendants were never brought into court in restraints. Again, this information indicates that the results presented here will underestimate the true effect of a meaningful ban on indiscriminate juvenile shackling. In 2016 (four years after the end of the NJCDA data), a juvenile court rule banning indiscriminate juvenile shackling

became effective in Tennessee.¹²⁹ A law to ban indiscriminate juvenile shackling was first proposed in the state in 2015 (three years after the end of the NJCDA data),¹³⁰ and the push for a change in the law by public defenders began in 2013 (a year after the end of the NJCDA data), making it unlikely that any anticipatory effects would exist in the Tennessee data.

D. National Juvenile Court Data Archive Data

This project utilizes data on juvenile court cases in North Carolina and Tennessee from the year 2004 to the year 2012. The data was generated by the North Carolina Department of Public Safety and the Tennessee Administrative Office of the Courts, and the data was made available for study by the National Juvenile Court Data Archive, which is maintained by the National Center for Juvenile Justice in Pittsburgh, Pennsylvania, and supported by a grant from the Office of Juvenile Justice and Delinquency Prevention, U.S. Department of Justice.¹³¹ Since the data are collected by state agencies for their own information and reporting needs, rather than based on a data collection mandate, there are differences in the information that Tennessee and North Carolina collect and in the states' methods of recording. This Part will describe how the data from the states were combined for analysis and provide summary statistics for each state.

The data from both states are analyzed at the case level, rather than the charge level or the individual level. Some juveniles are present in the data multiple times, and many cases involve more than one charge.¹³² The North Carolina data contain information on the county where the case was processed, and the Tennessee data contain information on the court where the case was

¹²⁹ TENN. R. JUV. PROC. 204 (2016).

¹³⁰ S.B. 608, 109th Sess. (Tenn. 2015).

¹³¹ NCJFCJ and NCJJ bear no responsibility for the analyses or interpretations presented herein.

¹³² The North Carolina data contain up to three charges per case. The Tennessee data contain up to five referral reasons (including dependency and neglect issues) per case. In both data sets, there may be more than one count of each charge/referral reason, allowing the total number of charges to exceed these limits.

processed. Since both states have one court with juvenile jurisdiction in each county (with the exception of a few counties that have two juvenile courts), the Tennessee court information is easily converted into county information to match the North Carolina data for analysis. The month and year assigned to a case for purposes of the double-difference and triple-difference analyses are the month and year of case disposal. Disposal month and year are not necessarily the same as case termination month and year. For example, a case is disposed when a court orders a term of probation, even though the case remains open pending successful completion of that probationary period.¹³³

Both datasets include cases where formal charges were filed as well as cases that were never formally petitioned, either because they were dropped during the intake process or resolved through diversion. Each case is coded as (1) dismissed, (2) transferred to adult court, or (3) resolved with a sentence. Cases are coded as having been transferred to adult court if any of the charges were transferred to adult court. Cases that were directly filed in adult court are not included in the data. Cases are coded as receiving a sentence if any of the charges received a sentence and none of the charges were transferred to adult court, and cases are coded as having been dismissed if all of the charges were dismissed or dropped. Sentences could be part of an informal diversion agreement or assigned by a judge after a finding of delinquency. Sentenced cases are grouped into four categories based on the most serious aspect of the sentence: detention, probation, treatment, and other. Treatment covers a range of medical and educational programs such as drug and alcohol counseling, mental health counseling, referrals to alternative school programs, and placement in group homes. Other sentences include fines, restitution, no contact orders, court defined curfews, public service, and similar court orders.

¹³³ I chose to use disposal month and year instead of termination month and year because, if the dates differ, the date of disposition will be closest to when the outcome variables of interest were decided.

The Tennessee data contain information on dependency and neglect cases as well as delinquency and status offense cases. If a case in the data only involves dependency and neglect issues, it is not included in the analyses presented here. Similarly, cases from both states that involve only minor infractions or traffic offenses (other than driving under the influence or driving while ability impaired) are excluded from the analyses.¹³⁴

Following Hockenberry and Puzanchera (2018), charges are grouped into five categories: crimes against persons, property crimes, drug crimes, crimes against public order, and status offenses. Crimes against persons include homicide, rape, robbery, assault, other sex offenses, kidnapping, harassment, and other similar offenses. Property crimes include arson, burglary, larceny, motor vehicle theft, vandalism, trespassing, fraud, and other similar offenses. Drug crimes include the possession, distribution, and manufacturing of illicit substances. Alcohol offenses are not included in drug crimes. Crimes against public order include obstruction of justice, disorderly conduct, weapons offenses, alcohol offenses (other than those which are status offenses), driving under the influence and driving while ability impaired, animal abuse, gambling, filing false police reports, and violations of wildlife regulations, among others. Finally, status offenses are those offenses that would not be a crime if the juvenile engaged in the same behavior as an adult. Common examples are underage drinking, possession of tobacco

¹³⁴ In the North Carolina data, these offenses include those that were coded as an infraction rather than a felony, misdemeanor, or status offense, those that were traffic related other than driving under the influence and driving while ability impaired, and a few other miscellaneous offenses such as littering. In the Tennessee data, where offenses are described at a more general level, these included traffic offenses other than driving under the influence and local ordinance violations. Following Hockenberry and Puzanchera (2018), minor infraction and traffic offense cases are excluded because they may be systematically different from other cases. If these cases are instead included (along with a dummy control variable equal to one when a minor infraction of traffic offense is charged), the results do not change substantially. The most notable difference is that the triple difference results indicate a significant increase in dismissals (by about 2 percentage points) and a decrease in sentencing (by about 1 percentage point), but only when the sample is not restricted by the age of the juvenile.

products, truancy, unruly behavior, and running away. About 10 percent of cases involve more than one type of crime.

Other case-level demographic information is available in addition to the information on crimes charged and case outcomes. For example, both states have information on the age of the juvenile at the time of referral,¹³⁵ the race and ethnicity of the juvenile,¹³⁶ the sex of the juvenile, and the length of the case.¹³⁷ Tennessee also provides information on school enrollment and living situation, and North Carolina also provides information on weapon use, whether there was a serious injury, and whether the offense was school-related. Only the Tennessee data contain information on pretrial detention.¹³⁸ For variables that are available in both states (and will be included in the double-difference and triple-difference analyses), missing values are replaced by regression-estimated values.¹³⁹ Table 2 contains case summary statistics for each state.

Table 2. Juvenile Case Summary Statistics

Variable	Tennessee	North Carolina
Number of Cases	553,641	240,796

¹³⁵ North Carolina also provides information on age at the time of the offense. In some delinquency cases, the reported age exceeds 15 (the oldest age for which the juvenile court would have jurisdiction), and in some status offense cases the reported age exceeds 17 (the oldest age for which the juvenile court would have jurisdiction). Since only cases with juvenile jurisdiction are included in the data, these must be the result of data entry errors; there are 259 such cases, and, while all crime types are represented, the majority of these cases involve person and property offenses. Age at offense is never included in the regression analyses (age at referral is used instead), and the results which restrict the sample to juveniles between the ages of 6 and 15 exclude these misreported observations, ensuring that they are not driving the results presented here.

¹³⁶ In Tennessee, race and ethnicity are reported separately, so that a juvenile may be coded as both black and Hispanic. In North Carolina, however, Hispanic is considered a race category. This means that a juvenile who is both black and Hispanic would be coded as multiracial in North Carolina. In order to make the data comparable, the Tennessee data on race and ethnicity are combined, resulting in those who are nonwhite and Hispanic being combined with those who are multiracial.

¹³⁷ In some instances (118 cases), case length is negative due to data entry errors. In these instances, case length is treated as a missing variable and replaced with the regression-estimated case length (see footnote 139).

¹³⁸ In the triple-difference analyses, pretrial detention is proxied for in North Carolina by whether the juvenile is charged with a felony offense.

¹³⁹ Regression-estimated values are calculated using a linear regression of all of the other case characteristics available in the state on the missing variable. The estimated regression equation is then used to predict a value of the missing variable based on the other case demographics available. This method is also used to replace missing values for pretrial detention in Tennessee and felony charges in North Carolina since these variables are used for triple-difference analyses.

Juvenile Demographics

Age at Referral	14.90	13.87
Age at Offense	—	13.77
Female	34.08%	30.33%
White, Non-Hispanic	59.16%	39.51%
White, Hispanic	1.31%	5.99%
Black, Non-Hispanic	37.45%	50.30%
Asian, Non-Hispanic	0.36%	0.52%
Native American, Non-Hispanic	0.08%	1.72%
Multi-Racial or Non-White, Hispanic	1.65%	1.95%
Living with Biological Parents	16.70%	—
Living with Remarried Parent	6.57%	—
Living with Single Parent	55.39%	—
Living with Relatives	8.51%	—
Living with Adopted Parents	0.69%	—
Placed Living Situation	3.94%	—
Living Independently	0.51%	—
Other Living Situation	1.47%	—
Enrolled in School	93.40 %	—
Enrolled in Special Education Program	8.68%	—

Case Demographics

Weapon Used	—	15.33%
Firearm/ Explosive Used	—	1.74%
Serious Injury	—	1.87%
School-Related Case	—	31.50%
Pretrial Detention	19.99%	—
Felony Charged	—	14.09%
Misdemeanor Charged	—	72.81%
Number of Initial Charges	1.24	1.94
Multiple Initial Charges	18.10%	73.24%
Charged with Crime Against Person	15.25%	34.49%
Charged with Property Crime	19.61%	32.03%
Charged with Drug or Alcohol Crime	7.56%	8.10%
Charged with Public Order Crime	29.74%	16.78%
Charged with Status Offense	36.07%	16.02%
Length of Case (In Months)	2.48	1.95

Case Outcomes

Dismissed	35.78%	55.49%
Transferred to Adult Court	0.54%	0.20%
Sentenced	53.40%	42.93%
Missing Case Outcome	10.27%	1.38%
Harshest Sentence - Detention	8.11%	23.22%
Harshest Sentence - Probation	39.75%	60.89%
Harshest Sentence - Treatment	17.41%	12.10%
Harshest Sentence - Other	34.73%	1.79%
Missing Sentence	—	2.01%

Notes: Observations are dropped from the Tennessee data if they do not involve any criminal charge, and observations are dropped from both states if they only involve traffic offenses or minor infractions. Observations from the North Carolina data are assigned the mean age at offense if the juvenile is over 17 at the time of the offense or over 15 at the time of the offense and is not charged with at least one status offense, and observations from the North Carolina data are recoded so that the age at offense is equal to the age at referral if the age at offense is greater than the age at referral. Case length is treated as missing if it is negative. Missing values for non-case outcome variables available in both states, pretrial detention, and felony charged are replaced with regression-estimated values, and missing values for variables only available in one state (other than pretrial detention and felony charged) are not included in this table. The sentencing percentages are calculated only for the sample of cases that result in a sentence. The charge categories are not mutually exclusive.

Unsurprisingly, the average age at referral is higher in Tennessee, where the statutory definition of a juvenile is more expansive. Tennessee also has more cases involving female juvenile defendants and fewer cases involving minority juvenile defendants than North Carolina. At least some of the racial disparity is explained by the fact that North Carolina is more diverse than Tennessee (North Carolina is 71 percent white-only and Tennessee is 79 percent white-only according to the U.S. Census Bureau (b, c)). In Tennessee, about 20 percent of cases involve pretrial detention, which is equal to the national average in 2008 (Puzzanchera, Adams, & Sickmund 2011). In North Carolina, only about 14 percent of cases involve a felony charge.

More cases in North Carolina involve multiple charges, cases take longer to resolve in Tennessee, and the distribution of charges between crime categories differs between the two states. In North Carolina, more cases involve crimes against persons and property and in Tennessee more cases involve status offenses and crimes against public order. Table 3 illustrates how the distribution of charges between crime categories in each state compares to the national average in 2008 (the midpoint in the data) (Puzzanchera, Adams, & Sickmund 2011). The national averages for property crimes and drug crimes lie above those of North Carolina and Tennessee, and the national average for status offenses is below those of North Carolina and Tennessee.

Table 3. Crime Category Distributions in North Carolina, Tennessee, and Nationwide (in 2008)

Crime Type	Tennessee	North Carolina	Nationwide in 2008
Charged with Crime Against Person	15.25%	34.49%	22%
Charged with Property Crime	19.61%	32.03%	34%
Charged with Drug or Alcohol Crime	7.56%	8.10%	10%
Charged with Public Order Crime	29.74%	16.78%	25%
Charged with Status Offense	36.07%	16.02%	12%

Notes: Observations are dropped from the Tennessee data if they do not involve any criminal charge, and observations are dropped from both states if they only involve traffic offenses or minor infractions. The charge categories are not mutually exclusive, except in the nationwide data, which only look at the most serious charge. The nationwide information is from Puzzanchera, Adams, and Sickmund (2011).

North Carolina also has more cases that result in a dismissal than Tennessee. Dismissals include voluntary dismissals by the prosecution and involuntary dismissals ordered by the judge. Nationwide in 2008, about 32 percent of cases were dismissed (Puzzanchera, Adams, & Sickmund 2011), a figure closer to the Tennessee figure than the North Carolina figure. Tennessee also has slightly more cases that are transferred to adult court, although transfer is fairly uncommon in both states. The national average for jurisdictional waivers to adult court in 2008 was 0.49 percent, a figure very similar to that of Tennessee (Puzzanchera, Adams, & Sickmund 2011).

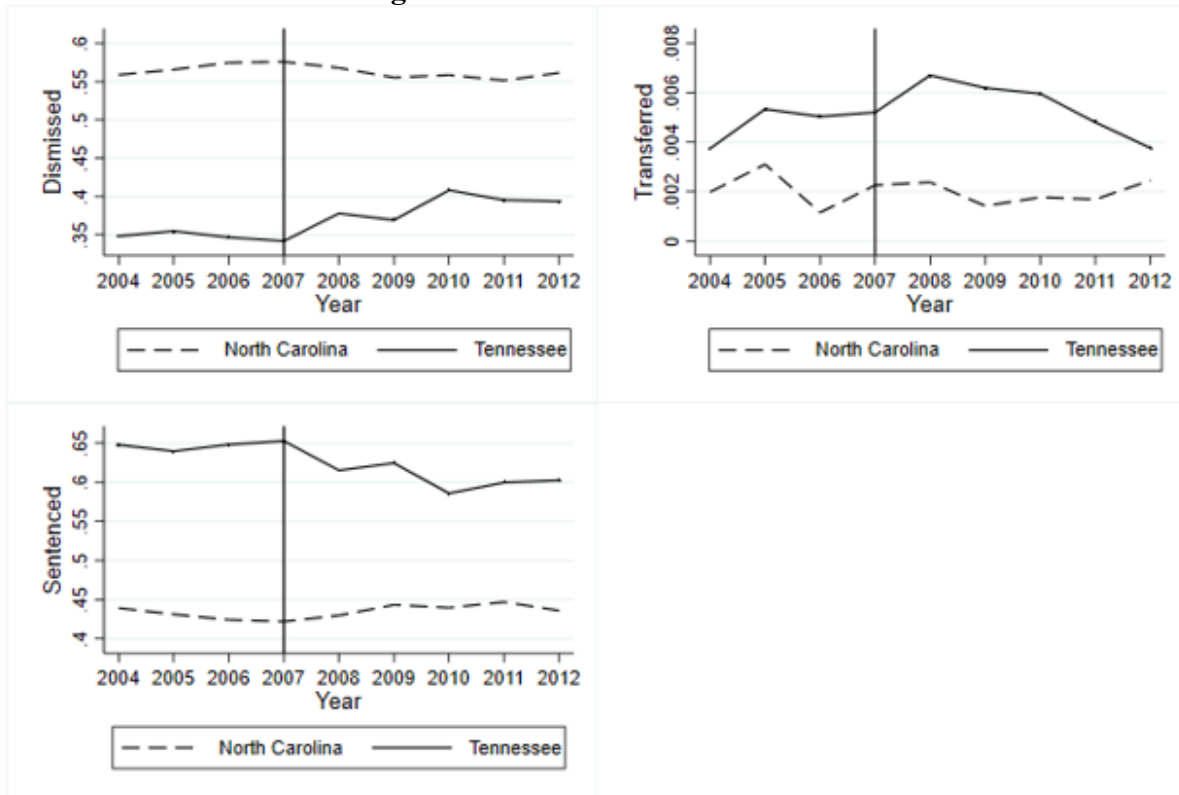
Surprisingly few of the differences between North Carolina and Tennessee are explained by differences in juvenile court jurisdiction: The differences remain even when the sample is restricted to juveniles between the ages of 6 and 15, inclusive, at the time of case referral. Appendix Table A3 contains case demographics for the age-restricted sample.

The differences in case characteristics correspond to differences in case resolutions. In North Carolina, where more cases that get to the sentencing stage involve crimes against persons and property, more cases result in detention and probation, and in Tennessee, where more cases

that get to the sentencing stage involve status offenses and crimes against public order, more cases result in treatment or some other sentence. Nationwide in 2008, about 9 percent of cases resulted in secure out-of-home placement, and about 21 percent of cases resulted in probation (Puzzanchera, Adams, & Sickmund 2011).

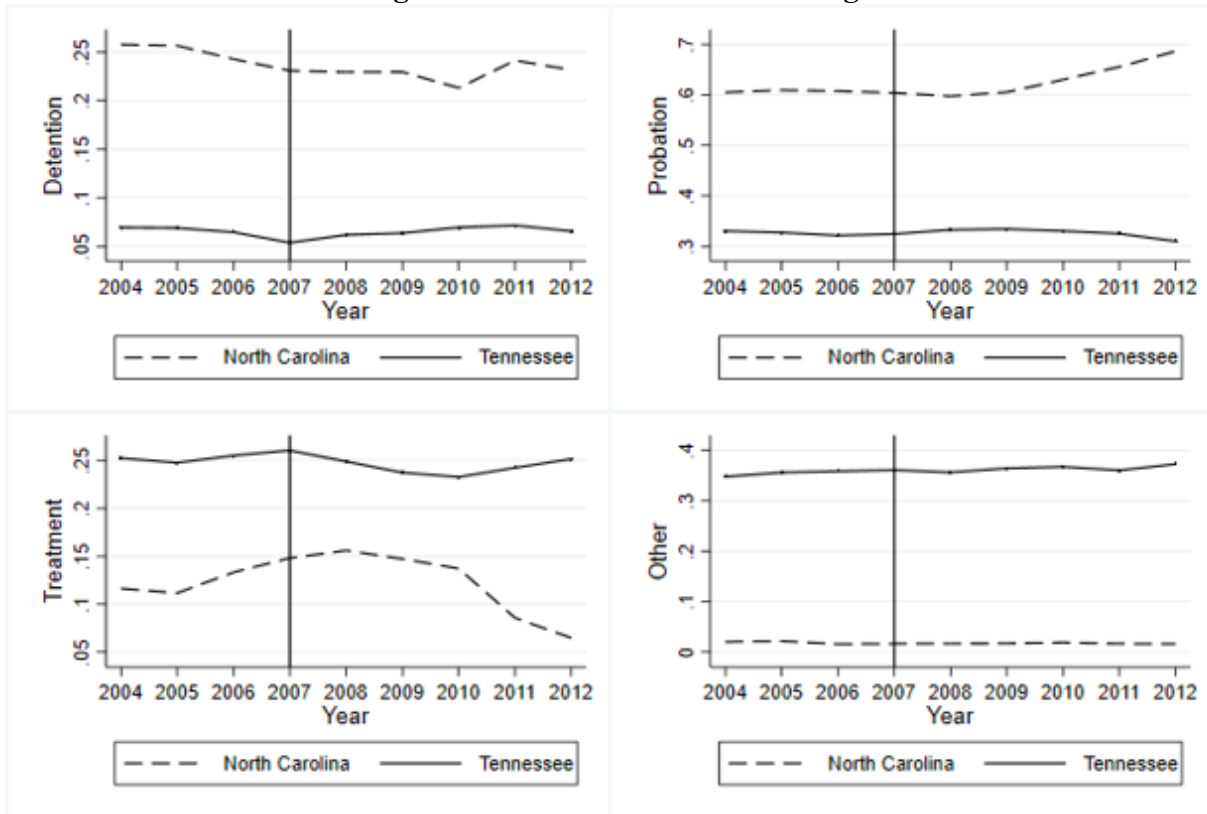
As discussed above, a critical assumption of the double-difference model is that the trends in outcomes would have been equivalent in North Carolina and Tennessee absent the passage of the ban on indiscriminate juvenile shackling in North Carolina. This is often referred to as the parallel trends assumption. Figures 2 and 3 illustrate the raw trends in case outcomes and sentencing in each state from 2004 to 2012. Although not perfect, the pre-2007 trends in North Carolina and Tennessee appear fairly parallel. The most notable break in the parallel trends is when North Carolina experienced a decrease in cases transferred to adult court between 2005 and 2006 and Tennessee did not. This is unlikely to be a major concern since the event of interest is in 2007 and the lines are relatively similar otherwise. The next Section will use the data and the methodologies just described to estimate the effects of North Carolina's ban on indiscriminate juvenile shackling on case outcomes.

Figure 2. Trends in Case Outcomes



Notes: Data are from the National Juvenile Court Data Archive. Cases are coded as dismissed if all of the charges were either dropped or dismissed, transferred if any of the charges were transferred to adult criminal court, and sentenced if any of the charges resulted in a sentence and none of the charges were transferred to adult criminal court.

Figure 3. Trends in Case Sentencing



Notes: Data are from the National Juvenile Court Data Archive. Cases are coded as sentenced if any of the charges resulted in a sentence and none of the charges were transferred to adult criminal court. Other sentences include minor sanctions such as a fine.

V. The Effect of Banning Indiscriminate Juvenile Shackling

A. Double-Difference Estimates

Table 4 contains the regression results for Equation 1, the double-difference estimation using the full sample of juveniles from both states described in detail above.

$$\begin{aligned}
 Outcome_{ict} &= \alpha + \beta_1 CaseControls_{ict} + \beta_2 County_c + \beta_3 Month_t \\
 &+ \beta_4 County_c * Trend_t + \beta_5 NorthCarolina_s * PostSept2007_t \\
 &+ \epsilon_{ict}
 \end{aligned} \tag{1}$$

The top row contains the coefficients of interest, representing the causal effect of North Carolina's ban on indiscriminate juvenile shackling on the various case outcomes. Figure 4

illustrates these double-difference coefficients for each of the case outcomes. The banning of indiscriminate juvenile shackling resulted in more favorable case outcomes for juveniles: Juveniles were 0.31 percentage points less likely to have their case transferred to adult court, a 65 percent decrease relative to the mean. Juveniles who received a sentence were also 4.10 percentage points less likely to receive probation, 4.99 percentage points more likely to receive treatment, and 0.78 percentage points less likely to receive some other sentence. The last result—that juveniles are less likely to receive some other sentence (such as a fine or community service)—is not necessarily a good result from the perspective of juvenile defendants. However, the size of this effect is dwarfed by the other, more positive results. Since all of the outcome variables are dummy variables equal to one if the case outcome occurred or the sentence type was assigned, significant changes in case outcomes represent shifts to or from other case outcomes and significant changes in sentence types represent shifts to or from other sentence types.

Importantly, these numbers represent an underestimation of the true effect of the ban for a number of reasons. There are some counties in North Carolina where the passage of the law did not change shackling behavior, either because juveniles were not being indiscriminately shackled before the law or because the law was not implemented as intended. There are also some counties in Tennessee that did not indiscriminately shackle juveniles during this time period, despite not having a ban in place. Finally, these numbers represent the average effect on all juveniles, including those who would not have been shackled regardless of the passage of the law because they were never detained pre-trial.

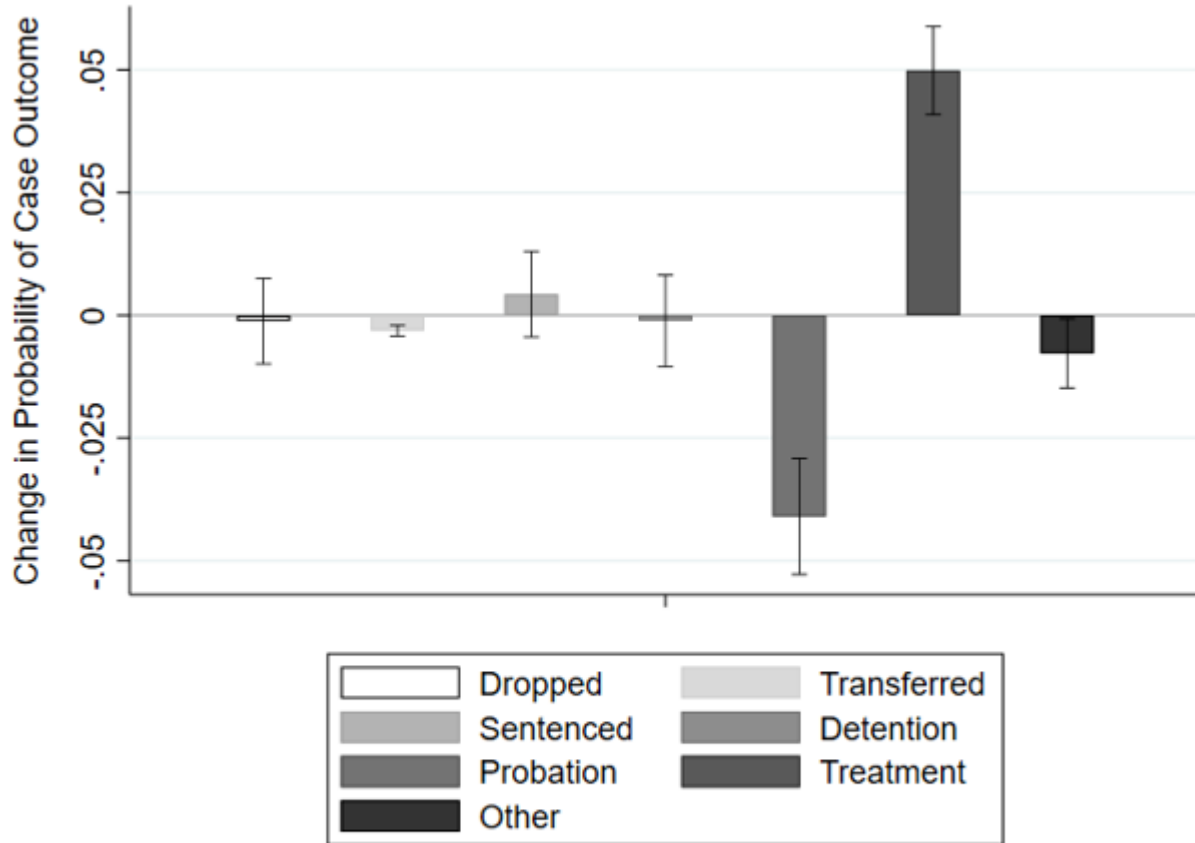
Table 4. Double-Difference Estimates of the Effect of North Carolina’s Indiscriminate Juvenile Shackling Ban on Case Outcomes

Variable	Dismissed	Transferred	Sentenced	Detention	Probation	Treatment	Other
NC*PostSept2007	-0.0012 (0.0045)	-0.0031*** (0.0006)	0.0043 (0.0045)	-0.0011 (0.0047)	-0.0410*** (0.0060)	0.0499*** (0.0046)	-0.0078** (0.0036)
Referral Age	-0.0168*** (0.0002)	0.0015*** (0.0000)	0.0152*** (0.0002)	0.0061*** (0.0002)	0.0094*** (0.0003)	-0.0035*** (0.0003)	-0.0120*** (0.0004)
Female	0.0517*** (0.0011)	-0.0049*** (0.0001)	-0.0468*** (0.0011)	-0.0245*** (0.0010)	0.0037** (0.0015)	0.0100*** (0.0012)	0.0108*** (0.0014)
Black, Non-Hispanic	-0.0246*** (0.0014)	0.0037*** (0.0002)	0.0210*** (0.0014)	0.0382*** (0.0014)	0.0136*** (0.0020)	-0.0243*** (0.0015)	-0.0275*** (0.0017)
Asian, Non-Hispanic	-0.0078 (0.0082)	0.0022* (0.0012)	0.0055 (0.0082)	0.0083 (0.0069)	0.0135 (0.0109)	-0.0210** (0.0091)	-0.0007 (0.0101)
Native American, Non-Hispanic	-0.0489*** (0.0087)	0.0022*** (0.0007)	0.0467*** (0.0087)	0.0143 (0.0109)	0.0506*** (0.0121)	-0.0351*** (0.0062)	-0.0298*** (0.0051)
White, Hispanic	-0.0059* (0.0034)	0.0026*** (0.0004)	0.0033 (0.0034)	0.0068** (0.0031)	0.0243*** (0.0044)	-0.0156*** (0.0032)	-0.0154*** (0.0033)
Non-White, Hispanic/Multiracial	-0.0444*** (0.0041)	0.0019*** (0.0005)	0.0425*** (0.0041)	0.0341*** (0.0038)	0.0174*** (0.0052)	-0.0219*** (0.0040)	-0.0297*** (0.0044)
Number of Charges	-0.0804*** (0.0010)	0.0053*** (0.0003)	0.0750*** (0.0011)	0.0252*** (0.0010)	0.0095*** (0.0014)	0.0133*** (0.0010)	-0.0480*** (0.0010)
Length of Case	0.0019*** (0.0001)	0.0001*** (0.0000)	-0.0019*** (0.0001)	0.0005*** (0.0001)	-0.0014*** (0.0002)	-0.0016*** (0.0001)	0.0025*** (0.0002)
Person Crime	-0.0010 (0.0020)	0.0090*** (0.0004)	-0.0080*** (0.0020)	0.0403*** (0.0020)	0.0259*** (0.0027)	-0.0429*** (0.0021)	-0.0233*** (0.0021)
Property Crime	-0.0679*** (0.0021)	0.0011** (0.0004)	0.0668*** (0.0021)	0.0217*** (0.0019)	0.0302*** (0.0027)	-0.0701*** (0.0021)	0.0182*** (0.0022)
Drug Crime	-0.0688*** (0.0024)	-0.0004 (0.0005)	0.0692*** (0.0024)	-0.0058*** (0.0021)	0.1065*** (0.0031)	0.0039 (0.0026)	-0.1047*** (0.0024)
Public Order Crime	0.0225*** (0.0019)	-0.0025*** (0.0004)	-0.0200*** (0.0019)	0.0598*** (0.0018)	0.0335*** (0.0025)	-0.0703*** (0.0019)	-0.0230*** (0.0020)
Status Offense	0.0650*** (0.0019)	-0.0030*** (0.0004)	-0.0620*** (0.0019)	-0.0120*** (0.0017)	-0.0429*** (0.0026)	-0.0277*** (0.0021)	0.0826*** (0.0021)
Constant	0.7178*** (0.0179)	-0.0284*** (0.0012)	0.3107*** (0.0179)	0.1037*** (0.0164)	0.3552*** (0.0230)	0.2925*** (0.0178)	0.2486*** (0.0179)
Observations	728,773	728,773	728,773	393,987	393,987	393,987	393,987
R-squared	0.2070	0.0180	0.2062	0.1982	0.2181	0.1261	0.2157

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

Notes: All estimations also include county and month fixed effects, dummy variables for whether the case demographics are missing, and county-specific linear time trends. Robust standard errors in parentheses.

Figure 4. Double-Difference Results



Notes: Data are from the National Juvenile Court Data Archive. Cases are coded as dismissed if all of the charges were either dropped or dismissed, transferred if any of the charges were transferred to adult criminal court, and sentenced if any of the charges resulted in a sentence and none of the charges were transferred to adult criminal court. Other sentences include minor sanctions such as a fine. Bars represent double-difference coefficients, and ninety-five percent confidence intervals are illustrated.

The results only change slightly when the sample is restricted to juveniles between the ages of 6 and 15, inclusive, to account for state differences in the legislative definition of a juvenile. There is still a significant decrease in the probability of receiving a sentence of probation (now 4.51 percentage points), a significant increase in the probability of receiving a sentence involving treatment (now 5.34 percentage points), and a significant decrease in the probability of receiving some other sentence (now 1.03 percentage points). The decrease in the probability of transfer to adult court also remains, although it is

only marginally significant: Juveniles are 0.08 percentage points less likely to have their case transferred, a 67 percent decrease relative to the mean. Table 5 contains these age-restricted regression results, and Figure 5 illustrates the double-difference coefficients for each of the case outcomes.

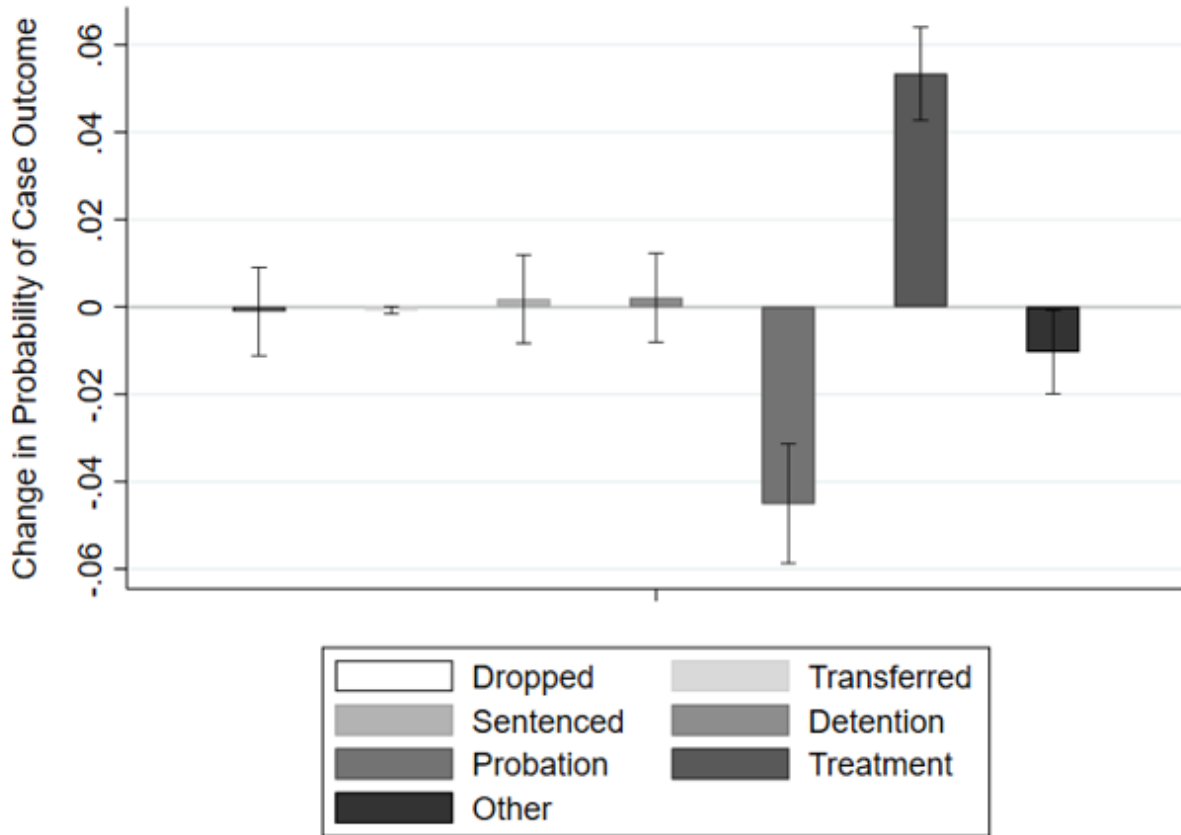
Table 5. Double-Difference Estimates for Juveniles Ages 6-15

Variable	Dismissed	Transferred	Sentenced	Detention	Probation	Treatment	Other
NC*PostSept2007	-0.0011 (0.0052)	-0.0008* (0.0004)	0.0018 (0.0052)	0.0021 (0.0052)	-0.0451*** (0.0070)	0.0534*** (0.0054)	-0.0103** (0.0049)
Referral Age	-0.0324*** (0.0004)	0.0004*** (0.0000)	0.0320*** (0.0004)	0.0105*** (0.0004)	0.0117*** (0.0006)	-0.0043*** (0.0005)	-0.0179*** (0.0006)
Female	0.0634*** (0.0014)	-0.0013*** (0.0001)	-0.0620*** (0.0014)	-0.0221*** (0.0013)	0.0102*** (0.0020)	0.0034** (0.0016)	0.0085*** (0.0017)
Black, Non-Hispanic	-0.0369*** (0.0018)	0.0013*** (0.0001)	0.0357*** (0.0018)	0.0308*** (0.0018)	0.0045* (0.0024)	-0.0137*** (0.0019)	-0.0217*** (0.0019)
Asian, Non-Hispanic	0.0023 (0.0105)	0.0033** (0.0014)	-0.0056 (0.0105)	-0.0020 (0.0093)	0.0269* (0.0143)	-0.0336*** (0.0108)	0.0087 (0.0114)
Native American, Non-Hispanic	-0.0615*** (0.0091)	-0.0000 (0.0005)	0.0616*** (0.0091)	0.0132 (0.0117)	0.0390*** (0.0125)	-0.0299*** (0.0064)	-0.0223*** (0.0046)
White, Hispanic	-0.0111*** (0.0038)	0.0013*** (0.0003)	0.0098** (0.0038)	0.0032 (0.0038)	0.0237*** (0.0051)	-0.0122*** (0.0038)	-0.0147*** (0.0033)
Non-White, Hispanic/Multiracial	-0.0619*** (0.0051)	0.0005 (0.0003)	0.0613*** (0.0051)	0.0261*** (0.0046)	0.0129** (0.0063)	-0.0131*** (0.0051)	-0.0258*** (0.0049)
Number of Charges	-0.0886*** (0.0013)	0.0017*** (0.0002)	0.0868*** (0.0013)	0.0260*** (0.0013)	-0.0018 (0.0017)	0.0099*** (0.0013)	-0.0341*** (0.0011)
Length of Case	-0.0020*** (0.0002)	0.0001*** (0.0000)	0.0020*** (0.0002)	0.0013*** (0.0002)	-0.0003 (0.0002)	-0.0018*** (0.0002)	0.0008*** (0.0002)
Person Crime	-0.0144*** (0.0027)	0.0016*** (0.0003)	0.0127*** (0.0027)	0.0331*** (0.0026)	0.0171*** (0.0034)	-0.0235*** (0.0027)	-0.0267*** (0.0023)
Property Crime	-0.0786*** (0.0028)	-0.0016*** (0.0003)	0.0802*** (0.0028)	0.0146*** (0.0027)	0.0207*** (0.0036)	-0.0419*** (0.0028)	0.0065** (0.0026)
Drug Crime	-0.0787*** (0.0034)	-0.0018*** (0.0003)	0.0805*** (0.0034)	0.0006 (0.0031)	0.0758*** (0.0043)	0.0094*** (0.0035)	-0.0858*** (0.0029)
Public Order Crime	0.0070*** (0.0025)	-0.0018*** (0.0003)	-0.0053** (0.0026)	0.0432*** (0.0025)	0.0255*** (0.0033)	-0.0482*** (0.0025)	-0.0206*** (0.0023)
Status Offense	0.0265*** (0.0027)	-0.0012*** (0.0002)	-0.0254*** (0.0027)	-0.0238*** (0.0025)	-0.0227*** (0.0035)	-0.0151*** (0.0028)	0.0616*** (0.0027)
Constant	0.9092*** (0.0252)	-0.0065*** (0.0005)	0.0973*** (0.0252)	0.0144 (0.0217)	0.4392*** (0.0316)	0.2809*** (0.0245)	0.2655*** (0.0237)
Observations	460,801	460,801	460,801	234,474	234,474	234,474	234,474
R-squared	0.1871	0.0063	0.1868	0.2552	0.2499	0.1498	0.2614

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

Notes: All estimations also include county and month fixed effects, dummy variables for whether the case demographics are missing, and county-specific linear time trends. Robust standard errors in parentheses.

Figure 5. Age-Restricted Double-Difference Results



Notes: Data are from the National Juvenile Court Data Archive. Cases are coded as dismissed if all of the charges were either dropped or dismissed, transferred if any of the charges were transferred to adult criminal court, and sentenced if any of the charges resulted in a sentence and none of the charges were transferred to adult criminal court. Other sentences include minor sanctions such as a fine. The sample is restricted to juveniles between the ages of 6 and 15, inclusive. Bars represent double-difference coefficients, and ninety-five percent confidence intervals are illustrated.

B. Triple-Difference Estimates

The double-difference estimates represent the causal effects of North Carolina’s ban on indiscriminate juvenile shackling as long as there was no other discrete change around the same time that affected case outcomes in either North Carolina or Tennessee. Even though background research revealed no such change, a triple-difference estimation can help increase confidence in the causal nature of the results, as well as make the results more precise, by estimating the

average effect on only juveniles most likely to be treated (those detained pretrial or charged with a felony) rather than the average effect on the entire population of juvenile defendants—some of whom would never have been shackled in the absence of a ban on indiscriminate shackling.

Table 6 contains the regression results for Equation 2, the triple-difference estimation described in detail above.

$$\begin{aligned}
 Outcome_{ict} &= \alpha + \beta_1 CaseControls_{ict} + \beta_2 County_c + \beta_3 Month_t \\
 &+ \beta_4 PreTrialDetention_{ict} + \beta_5 County_c * Month_t + \beta_6 County_c \\
 &* PretrialDetention_{ict} + \beta_7 Month_t * PretrialDetention_{ict} \\
 &+ \beta_8 NorthCarolina_s * PostSept2007_t * PretrialDetention_{ict} + \epsilon_{ict}
 \end{aligned} \tag{2}$$

Again, the top row contains the coefficients of interest, representing the causal effect of North Carolina’s ban on indiscriminate juvenile shackling on the case outcomes of juveniles most likely to be impacted by the ban, and Figure 6 illustrates these triple-difference coefficients for each case outcome. Again, the results illustrate that the banning of indiscriminate juvenile shackling resulted in more favorable case outcomes for juveniles: Juveniles are 0.39 percentage points less likely to have their cases transferred to adult court (an 81 percent decrease relative to the mean). Additionally, juveniles who receive a sentence are 3.42 percentage points less likely to receive detention and 4.04 percentage points more likely to receive treatment.

Table 6. Triple-Difference Estimates of the Effect of North Carolina’s Indiscriminate Juvenile Shackling Ban on Case Outcomes

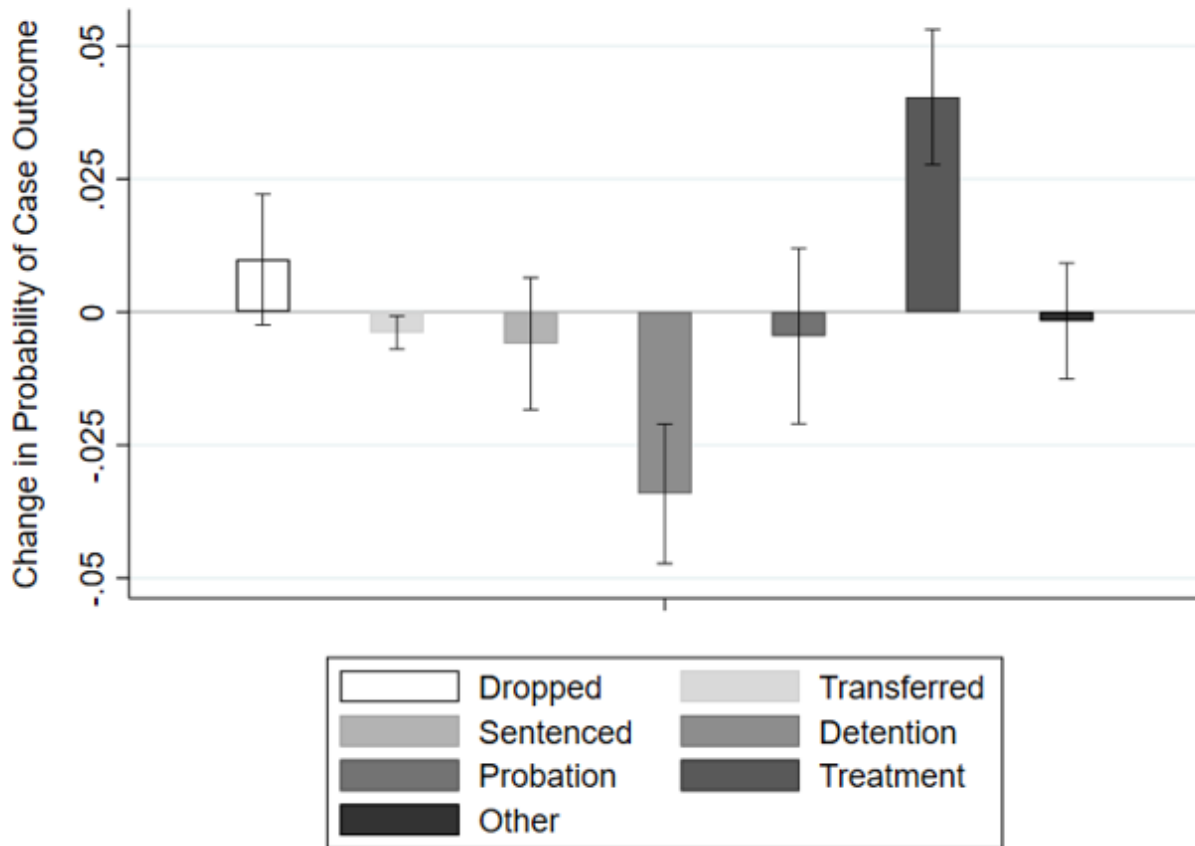
Variable	Dismissed	Transferred	Sentenced	Detention	Probation	Treatment	Other
NC*PostSept2007*	0.0099	-0.0039**	-0.0060	-0.0342***	-0.0045	0.0404***	-0.0017
Pretrial Detention	(0.0063)	(0.0016)	(0.0063)	(0.0067)	(0.0084)	(0.0065)	(0.0056)
Referral Age	-0.0154***	0.0015***	0.0139***	0.0051***	0.0075***	-0.0036***	-0.0090***
	(0.0002)	(0.0000)	(0.0002)	(0.0002)	(0.0004)	(0.0003)	(0.0004)
Female	0.0394***	-0.0043***	-0.0351***	-0.0207***	0.0005	0.0112***	0.0090***
	(0.0011)	(0.0001)	(0.0011)	(0.0010)	(0.0016)	(0.0013)	(0.0015)
Black, Non-Hispanic	-0.0168***	0.0031***	0.0137***	0.0361***	0.0125***	-0.0248***	-0.0238***
	(0.0014)	(0.0002)	(0.0014)	(0.0014)	(0.0020)	(0.0015)	(0.0017)
Asian, Non-	-0.0036	0.0027**	0.0008	0.0077	0.0051	-0.0155*	0.0027

Hispanic	(0.0080)	(0.0012)	(0.0080)	(0.0068)	(0.0110)	(0.0091)	(0.0104)
Native American, Non-Hispanic	-0.0448*** (0.0086)	0.0018** (0.0008)	0.0430*** (0.0086)	0.0044 (0.0113)	0.0489*** (0.0125)	-0.0299*** (0.0067)	-0.0234*** (0.0053)
White, Hispanic	-0.0016 (0.0033)	0.0024*** (0.0004)	-0.0008 (0.0033)	0.0066** (0.0031)	0.0203*** (0.0045)	-0.0150*** (0.0033)	-0.0119*** (0.0034)
Non-White, Hispanic/Multiracial	-0.0360*** (0.0040)	0.0017*** (0.0005)	0.0343*** (0.0041)	0.0325*** (0.0037)	0.0117** (0.0053)	-0.0194*** (0.0040)	-0.0248*** (0.0045)
Number of Charges	-0.0619*** (0.0011)	0.0056*** (0.0003)	0.0563*** (0.0011)	0.0180*** (0.0011)	0.0205*** (0.0015)	0.0107*** (0.0011)	-0.0492*** (0.0012)
Length of Case	0.0023*** (0.0001)	-0.0000 (0.0000)	-0.0023*** (0.0001)	0.0005*** (0.0001)	-0.0013*** (0.0002)	-0.0016*** (0.0001)	0.0024*** (0.0002)
Person Crime	-0.0199*** (0.0020)	0.0080*** (0.0005)	0.0119*** (0.0021)	0.0361*** (0.0020)	0.0042 (0.0028)	-0.0359*** (0.0022)	-0.0043* (0.0022)
Property Crime	-0.0540*** (0.0021)	-0.0011** (0.0005)	0.0551*** (0.0021)	0.0145*** (0.0019)	0.0214*** (0.0028)	-0.0656*** (0.0022)	0.0297*** (0.0023)
Drug Crime	-0.0675*** (0.0024)	-0.0021*** (0.0005)	0.0696*** (0.0024)	-0.0092*** (0.0021)	0.0918*** (0.0032)	0.0086*** (0.0027)	-0.0912*** (0.0026)
Public Order Crime	0.0106*** (0.0019)	-0.0041*** (0.0004)	-0.0065*** (0.0019)	0.0578*** (0.0018)	0.0127*** (0.0026)	-0.0621*** (0.0020)	-0.0084*** (0.0021)
Status Offense	0.0403*** (0.0020)	-0.0035*** (0.0004)	-0.0368*** (0.0020)	-0.0061*** (0.0018)	-0.0630*** (0.0027)	-0.0228*** (0.0022)	0.0920*** (0.0023)
Constant	0.7651*** (0.0040)	-0.0245*** (0.0006)	0.2594*** (0.0040)	-0.0103*** (0.0035)	0.3113*** (0.0058)	0.2518*** (0.0047)	0.4471*** (0.0059)
Observations	727,935	727,935	727,935	392,517	392,517	392,517	392,517
R-squared	0.2820	0.0546	0.2795	0.2872	0.2971	0.2139	0.2848

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

Notes: All estimations also include county and month fixed effects, county-by-month fixed effects, month-by-pretrial detention fixed effects, county-by-pretrial detention fixed effects, and dummy variables for whether the case demographics are missing. Robust standard errors in parentheses.

Figure 6. Triple-Difference Results



Notes: Data are from the National Juvenile Court Data Archive. Cases are coded as dismissed if all of the charges were either dropped or dismissed, transferred if any of the charges were transferred to adult criminal court, and sentenced if any of the charges resulted in a sentence and none of the charges were transferred to adult criminal court. Other sentences include minor sanctions such as a fine. Bars represent triple-difference coefficients, and ninety-five percent confidence intervals are illustrated.

The results change slightly when the sample is restricted to juveniles between the ages of 6 and 15, inclusive, to account for legislative differences in the definition of a juvenile. Juvenile defendants are marginally significantly less likely to receive a sentence (by 1.27 percentage points), and there is no longer a significant effect on the likelihood of transfer to adult court, suggesting that this result is driven by the older juveniles present in the Tennessee data. Juveniles who are sentenced are still less likely to receive detention (by 3.63 percentage points) and more likely to receive treatment (by 4.13 percentage points), but they are also significantly less likely

to receive some other sentence (by 1.88 percentage points). Again, the decrease in the probability of some other (relatively minor) sentence may not be a good outcome from the perspective of juvenile defendants, but the size of the effect is small relative to the decrease in the probability of detention and the increase in the probability of treatment. Table 7 contains the regression results, restricting the sample to juveniles between the ages of 6 and 15, inclusive, at the time of case referral, and Figure 7 depicts the triple-difference coefficients for each case outcome.

Table 7. Triple-Difference Estimates for Juveniles Ages 6-15

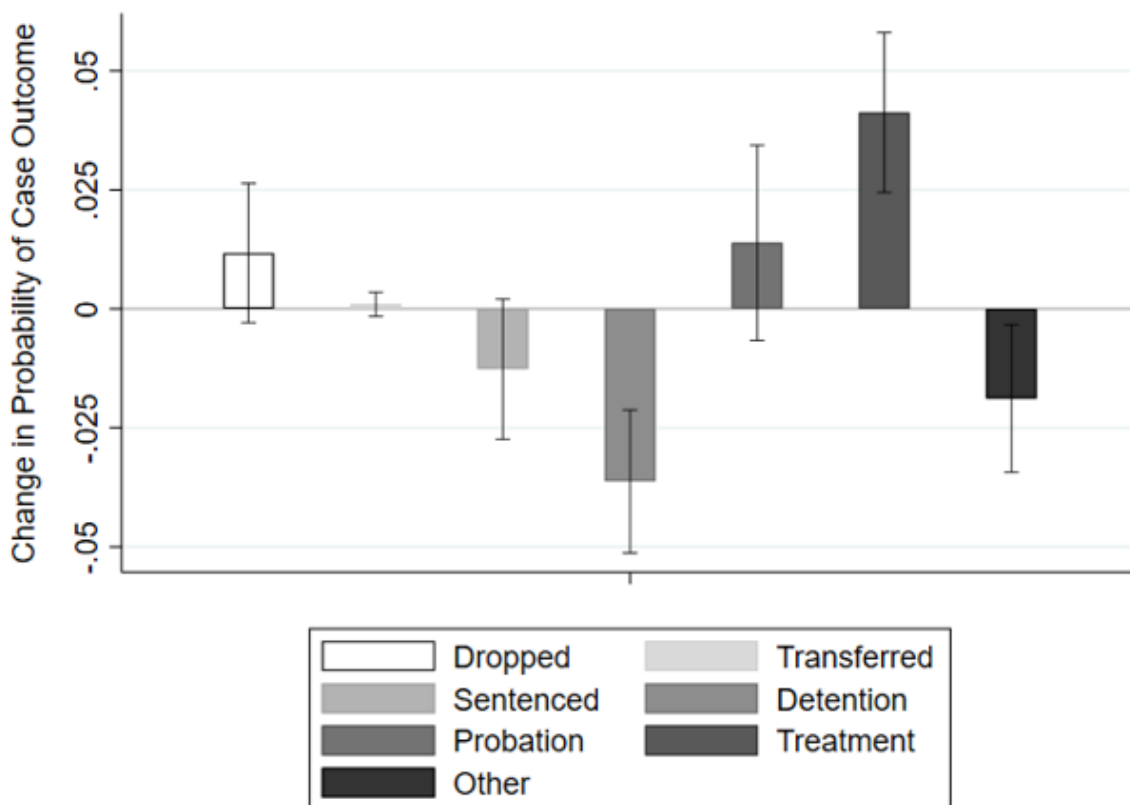
Variable	Dismissed	Transferred	Sentenced	Detention	Probation	Treatment	Other
NC*PostSept2007*	0.0117	0.0010	-0.0127*	-0.0363***	0.0139	0.0413***	-0.0188**
Pretrial Detention	(0.0075)	(0.0013)	(0.0075)	(0.0077)	(0.0105)	(0.0086)	(0.0079)
Referral Age	-0.0298***	0.0004***	0.0294***	0.0092***	0.0091***	-0.0040***	-0.0144***
	(0.0004)	(0.0000)	(0.0004)	(0.0004)	(0.0006)	(0.0005)	(0.0006)
Female	0.0481***	-0.0008***	-0.0473***	-0.0182***	0.0058***	0.0043***	0.0081***
	(0.0014)	(0.0001)	(0.0014)	(0.0014)	(0.0021)	(0.0017)	(0.0018)
Black, Non-Hispanic	-0.0302***	0.0011***	0.0292***	0.0297***	0.0058**	-0.0159***	-0.0196***
	(0.0018)	(0.0001)	(0.0018)	(0.0018)	(0.0025)	(0.0020)	(0.0020)
Asian, Non-Hispanic	0.0041	0.0036**	-0.0076	-0.0025	0.0250*	-0.0330***	0.0105
	(0.0102)	(0.0014)	(0.0103)	(0.0094)	(0.0145)	(0.0110)	(0.0118)
Native American, Non-Hispanic	-0.0585***	0.0000	0.0585***	0.0052	0.0405***	-0.0264***	-0.0193***
	(0.0091)	(0.0006)	(0.0091)	(0.0124)	(0.0131)	(0.0070)	(0.0048)
White, Hispanic	-0.0065*	0.0012***	0.0052	0.0043	0.0219***	-0.0132***	-0.0131***
	(0.0038)	(0.0003)	(0.0038)	(0.0039)	(0.0053)	(0.0038)	(0.0035)
Non-White, Hispanic/Multiracial	-0.0527***	0.0004	0.0523***	0.0248***	0.0086	-0.0113**	-0.0221***
	(0.0051)	(0.0004)	(0.0051)	(0.0046)	(0.0066)	(0.0052)	(0.0052)
Number of Charges	-0.0625***	0.0011***	0.0614***	0.0201***	0.0119***	0.0027*	-0.0347***
	(0.0015)	(0.0002)	(0.0015)	(0.0015)	(0.0019)	(0.0014)	(0.0013)
Length of Case	-0.0013***	0.0000**	0.0013***	0.0013***	-0.0004	-0.0017***	0.0008***
	(0.0002)	(0.0000)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Person Crime	-0.0448***	0.0025***	0.0423***	0.0326***	-0.0087**	-0.0146***	-0.0094***
	(0.0027)	(0.0003)	(0.0027)	(0.0027)	(0.0036)	(0.0028)	(0.0025)
Property Crime	-0.0657***	-0.0022***	0.0679***	0.0077***	0.0091**	-0.0329***	0.0161***
	(0.0029)	(0.0003)	(0.0029)	(0.0028)	(0.0038)	(0.0029)	(0.0027)
Drug Crime	-0.0774***	-0.0018***	0.0792***	-0.0032	0.0610***	0.0160***	-0.0738***
	(0.0034)	(0.0003)	(0.0034)	(0.0032)	(0.0045)	(0.0036)	(0.0031)
Public Order Crime	-0.0169***	-0.0010***	0.0180***	0.0441***	0.0007	-0.0380***	-0.0068***
	(0.0026)	(0.0003)	(0.0026)	(0.0026)	(0.0035)	(0.0027)	(0.0025)
Status Offense	-0.0077***	-0.0003	0.0081***	-0.0179***	-0.0449***	-0.0075**	0.0702***
	(0.0028)	(0.0002)	(0.0028)	(0.0026)	(0.0037)	(0.0029)	(0.0029)
Constant	1.0304***	-0.0055***	-0.0249***	-0.0399***	0.3443***	0.2420***	0.4537***

	(0.0057)	(0.0004)	(0.0057)	(0.0058)	(0.0090)	(0.0075)	(0.0084)
Observations	459,504	459,504	459,504	232,415	232,415	232,415	232,415
R-squared	0.2821	0.0725	0.2803	0.3613	0.3506	0.2627	0.3517

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

Notes: All estimations also include county and month fixed effects, county-by-month fixed effects, month-by-pretrial detention fixed effects, county-by-pretrial detention fixed effects, and dummy variables for whether the case demographics are missing. Robust standard errors in parentheses.

Figure 7. Age-Restricted Triple-Difference Results



Notes: Data are from the National Juvenile Court Data Archive. Cases are coded as dismissed if all of the charges were either dropped or dismissed, transferred if any of the charges were transferred to adult criminal court, and sentenced if any of the charges resulted in a sentence and none of the charges were transferred to adult criminal court. Other sentences include minor sanctions such as a fine. The sample is restricted to juveniles between the ages of 6 and 15, inclusive. Bars represent triple-difference coefficients, and ninety-five percent confidence intervals are illustrated.

VI. Heterogeneous Effects Across Crimes and Juveniles

The results presented above illustrate that North Carolina’s ban on indiscriminate juvenile shackling improved case outcomes for juveniles. It is possible, however, that the

majority of these benefits are concentrated among certain juveniles or certain types of cases. The results from Chapter Two revealed that shackling has very little effect on trial outcomes in the context of a nonviolent property crime in adult court. If this finding carries over to the juvenile context, it would suggest that the majority of the effects are concentrated among non-property crime cases. Moreover, the effects in different cases for different types of defendants could be counteracting. For example, seeing a 15 year old male defendant accused of assault in shackles may lead decisionmakers to make assumptions about dangerousness—resulting in harsher case outcomes—while seeing a 10 year old female defendant accused of truancy in shackles may engender feelings of sympathy in decisionmakers—resulting in more lenient case outcomes.

In order to examine whether the effects of the ban on indiscriminate juvenile shackling are heterogeneous across juveniles and case types, Equation 2 is estimated for (1) female defendants, (2) male defendants, (3) white defendants, (4) black defendants, (5) defendants between the ages of 6 and 13, inclusive, (6) defendants 14 or 15 years old, (7) defendants charged with a crime against a person, (8) defendants charged with a property crime, (9) defendants charged with a drug crime, (10) defendants charged with a crime against public order, and (11) defendants charged with a status offense. In order to maintain adequate sample sizes, juveniles of all ages are included (except where heterogeneity by age is being examined). The triple-difference coefficients of interest are presented in Table 8.

Table 8. Triple-Difference Estimates for Various Subsets of Cases

Variable	Dismissed	Transferred	Sentenced	Detention	Probation	Treatment	Other
Male	0.0143** (0.0072)	-0.0034* (0.0019)	-0.0109 (0.0072)	-0.0352*** (0.0077)	0.0039 (0.0097)	0.0239*** (0.0075)	0.0075 (0.0066)
Female	0.0090 (0.0180)	-0.0018 (0.0026)	-0.0072 (0.0180)	-0.0246 (0.0178)	-0.0049 (0.0220)	0.0536*** (0.0161)	-0.0241* (0.0127)
White	0.0110 (0.0103)	-0.0053*** (0.0016)	-0.0057 (0.0104)	-0.0242** (0.0099)	-0.0265** (0.0132)	0.0256** (0.0102)	0.0251*** (0.0077)

Black	0.0103 (0.0088)	-0.0024 (0.0025)	-0.0079 (0.0089)	-0.0682*** (0.0102)	0.0401*** (0.0127)	0.0689*** (0.0104)	-0.0408*** (0.0097)
Ages 6–13	-0.0198 (0.0150)	0.0009 (0.0008)	0.0188 (0.0150)	-0.0283* (0.0147)	-0.0316 (0.0213)	0.0422** (0.0182)	0.0177 (0.0158)
Ages 14–15	0.0262*** (0.0090)	0.0008 (0.0017)	-0.0271*** (0.0090)	-0.0407*** (0.0093)	0.0392*** (0.0126)	0.0493*** (0.0102)	-0.0478*** (0.0096)
Person Crime	0.0015 (0.0142)	0.0040 (0.0066)	-0.0054 (0.0149)	-0.0853*** (0.0174)	0.0326 (0.0207)	0.0419*** (0.0158)	0.0109 (0.0138)
Property Crime	-0.0247** (0.0108)	-0.0083*** (0.0025)	0.0330*** (0.0109)	-0.0439*** (0.0115)	0.0451*** (0.0149)	0.0568*** (0.0116)	-0.0580*** (0.0114)
Drug Crime	-0.0009 (0.0225)	0.0047 (0.0044)	-0.0038 (0.0226)	-0.0522** (0.0239)	0.0797** (0.0313)	-0.0386 (0.0243)	0.0111 (0.0186)
Public Order Crime	-0.0353* (0.0206)	-0.0021 (0.0034)	0.0374* (0.0208)	-0.0148 (0.0238)	-0.0364 (0.0273)	0.0067 (0.0192)	0.0445*** (0.0115)
Status Offense	-0.0759 (0.0871)	-0.0009 (0.0017)	0.0768 (0.0870)	-0.4417*** (0.1101)	0.3790*** (0.1389)	0.1183* (0.0704)	-0.0556* (0.0338)

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

Notes: All estimations also include county and month fixed effects, county-by-month fixed effects, month-by-pretrial detention fixed effects, county-by-pretrial detention fixed effects, controls for the same case demographics presented in Table 5, and dummy variables for whether the case demographics are missing. Robust standard errors in parentheses.

The results suggest that older males benefit the most from bans on indiscriminate juvenile shackling. Female juveniles are significantly more likely to receive a treatment-based sentence (by almost 6 percentage points), but the only other marginally significant result for female juveniles is that they are significantly less likely to receive some other (relatively minor) sentence. Nevertheless, it is important to note that the sample of female juvenile defendants is smaller than the sample of male juvenile defendants, which could be responsible for some of the lack of significance. Similarly, juveniles ages 6–13 are significantly more likely to receive a treatment-based sentence, and marginally significantly less likely to receive a detention sentence, but these effects are smaller in magnitude than those for older juveniles, who also experience additional significant benefits.

The race-based comparison is more complex. White juveniles are significantly less likely to have their cases transferred to adult court, significantly less likely to be sentenced to detention

or probation, and significantly more likely to be sentenced to treatment or some other sanction (like a fine or community service). All of the sentencing effects are between 2 and 3 percentage points. The effects on the likelihoods of receiving a detention sentence and a treatment-based sentence are much larger for black juveniles—almost seven percentage points. However, black juveniles are also significantly less likely to receive some other (relatively minor) sentence and significantly more likely to be sentenced to probation (each by about 4 percentage points). They also do not experience a significant decrease in the probability of transfer to adult court.

The effects also differ by crime type. The ban resulted in a decrease in the probability of receiving a sentence of detention and an increase in the probability of receiving a treatment-based sentence in cases involving crimes against people. Similarly, in drug crime cases, the ban resulted in a decrease in the probability of receiving a sentence of detention and an increase in the probability of receiving a sentence of probation. In property crime cases and cases involving crimes against public order, however, the effects are not universally positive for juvenile defendants. For property crime cases, the probability of a case being transferred to adult court decreased, the probability of receiving a sentence of detention decreased, and the probability of receiving a treatment-based sentence increased. In contrast, the probability of a case being dismissed decreased, the probability of a case resulting in a sentence increased, and the probability of receiving some other sentence (such as a fine) decreased. These results, in combination with those from Chapter Two, suggest that there is something unique about property crime cases in terms of the effects of defendant appearance. In cases involving a crime against public order, the probability of receiving some other sentence (such as a fine) went up, but the probability of dismissal went down, and the probability of receiving some kind of sentence went up (although both of these effects are only significant at the 10 percent level).

Finally, the effect of the ban on outcomes of cases involving status offenses is enormous. The probability of receiving a sentence involving detention decreased by 44.17 percentage points, the probability of receiving a sentence of probation increased by 37.9 percentage points, the probability of receiving a treatment-based sentence increased by 11.83 percentage points, and the probability of receiving some other sentence decreased by 5.56 percentage points. The relative size of these coefficients may be based on the relatively small number of cases (118) in North Carolina that involve a status offense as well as a felony offense (the variable used as a proxy for pretrial detention in North Carolina). Nevertheless, these results suggest that shackling in status offense cases—arguably the cases where shackling is most likely to be extraneous—is particularly harmful for juvenile defendants.

VII. Discussion and Conclusion

The results presented above illustrate that North Carolina's ban on indiscriminate juvenile shackling resulted in more favorable case outcomes for juveniles. Specifically, the triple-difference results show a decrease in the probability of detention and an increase in the probability of treatment among juveniles who receive a sentence. These effects are large—representing changes of three to four percentage points. The size and significance of the effects are particularly striking, given the fact that the estimates represent lower bounds on the actual size of the effects. Some counties in North Carolina may not have been indiscriminately shackling juvenile defendants even prior to the passage of the law, and some counties in Tennessee were not indiscriminately shackling juveniles even though it was legal.

There are at least two reasons why a ban on indiscriminate juvenile shackling could lead to better outcomes for juveniles. First, decisionmakers may be influenced by the appearance of shackled juvenile defendants. If decisionmakers associate shackling with dangerousness, for

example, then a ban on indiscriminate juvenile shackling would lead to more favorable case outcomes for juveniles. In order to cause the results presented here, the shackles themselves (and not the fact that they signal pretrial detention, for example) must bias the decisionmaker. North Carolina's legislative ban on indiscriminate juvenile shackling did nothing to change decisionmakers' access to other information (such as that related to pretrial detention); the law only narrowed the circumstances under which a judge could allow shackling, necessarily reducing the percentage of defendants actually shackled.

The results could also be caused by changes in juvenile defendant behavior brought about by the banning of indiscriminate juvenile shackling. For example, if juveniles feel less ashamed and are better able to communicate with their attorneys and the judge when they are not shackled, that behavior could lead to more favorable case outcomes. Exploring which of these mechanisms is primarily responsible for the results presented here remains an important issue for future research.

The results are primarily concentrated among older, male juveniles who are disproportionately represented in the juvenile justice system, and the effects on the probability of detention and treatment-based sentences are largest (about 7 percentage points) for black juveniles. This heterogeneity either means that older, black, male juveniles' behavior is more affected by shackles or that decisionmakers are more biased against older, black, male juveniles when they see them in shackles. It also suggests that a nationwide ban on indiscriminate juvenile shackling may reduce some of the race, gender, and age-based disparities currently present in the juvenile justice system.

The effects also differ based on the crimes charged. While outcomes universally improve in cases involving drug crimes, crimes against persons, and status offenses, the results are mixed

in cases involving property crimes and crimes against public order. Sentencing outcomes improve for juveniles that receive a sentence, but fewer cases are dismissed and more cases result in a sentence. If the ban is primarily affecting case outcomes by changing juvenile behavior, this would indicate that juvenile defendants charged with drug offenses, status offenses, and crimes against persons are more impacted by the presence of shackles (and therefore gain the most from the passage of the ban). If instead the ban is primarily affecting case outcomes by reducing appearance bias, this would indicate that decisionmakers are more biased against juvenile defendants charged with drug offenses, status offenses, and crimes against persons based on the presence of shackles.

The results presented here stand in contrast to the results of Chapter Two, which found a very limited effect of shackles in the context of an adult criminal trial for a nonviolent property crime. There are a few possible explanations. First, shackling may simply have a greater effect in the juvenile context. This could be because the sight of a shackled juvenile defendant serves as a greater signal of dangerousness than the sight of a shackled adult defendant or because juvenile defendants' behavior is more affected by shackling than adult defendants' behavior.

Alternatively, and perhaps more likely, the effects of shackling may be parallel in the adult and juvenile contexts, and the difference in results may be explained by differences in the scopes of the studies. This could be the case if the effect of shackling in the adult context—like in the juvenile context—is heterogeneous based on the crime charged. Chapter Two focuses exclusively on nonviolent property crime cases, which the results in this Chapter suggest may be one of a few types of cases in which shackling does not influence case outcomes in a clear-cut direction. This could also be the case if shackling only influences case outcomes outside the structured environment of a trial (for example, through dismissals) or if shackling only

influences decisionmakers choice between types of punishment and not the length of punishment given a sentence of incarceration. Finally, this could be the case if shackling influences case outcomes by influencing defendant behavior, as Chapter Two focuses exclusively on the appearance effect of shackling.

Whatever the cause of the results and the difference between the findings presented here and in Chapter Two, the results from this Chapter suggest that shackling is inappropriately influencing juvenile case outcomes, providing strong empirical support for a national ban on indiscriminate juvenile shackling, either through an extension of the Due Process rights recognized in *Deck v. Missouri* or through legislative or regulatory action.¹⁴⁰

While the holding of *Deck* does not directly apply to the context examined here—these are juvenile proceedings before a judge rather than a jury and many of the results relate to sentencing decisions rather than determinations of guilt—some of the concerns expressed by the Court in *Deck* may have similar force here. As in *Deck*, the government is taking action to restrain defendants in a way that could undermine the presumption of innocence at the guilt determination phase, interfere with communication between defendant and counsel, and undermine the dignity and rehabilitative goals of the juvenile system.¹⁴¹ The results in this Chapter may indicate that shackling undermines the presumption of innocence—in some specifications the ban significantly reduced the odds that a juvenile received any sentence, which is the case outcome most directly related to the determination of guilt. More definitively, the results presented here illustrate that indiscriminate shackling undermines the rehabilitative goals of the juvenile system—juveniles had a significantly higher probability of being sentenced to detention and a significantly lower probability of receiving a treatment-based sentence when

¹⁴⁰ 544 U.S. 622 (2005).

¹⁴¹ *Id.*

indiscriminate juvenile shackling was permitted. The results also support an argument that Due Process protections are necessary to fundamental fairness, as required by the Court in *In Re Gault*.¹⁴² State intervention through shackling results in harsher case outcomes specifically for juveniles detained pretrial. This means that, on average, indiscriminate juvenile shackling results in juveniles who are detained pretrial being worse off in terms of their case outcomes than they would be absent pretrial detention.

Even if the Court were to hold that indiscriminate juvenile shackling is not unconstitutional, the results in this Chapter provide sufficient justification for more legislative or regulatory state-wide bans. In considering whether to adopt these bans and how to design them, however, policymakers should consider both the costs and the benefits.¹⁴³ A complete cost-benefit analysis is beyond the scope of this paper and will likely be state-specific, but a generalized taxonomy of potential costs and benefits may be illustrative.

First, this paper illustrates three distinct benefits stemming from a state-wide ban on indiscriminate juvenile shackling: (1) greater equality and fairness, (2) increased opportunity for rehabilitation, and (3) either a decrease in the impact of appearance bias or a change in juvenile behavior leading to more lenient case outcomes. As discussed briefly above, the results represent an increase in equality and fairness because only juveniles who are subject to pretrial detention will experience the harsher case outcomes associated with indiscriminate juvenile shackling. Increased fairness is a benefit of the ban in and of itself, and it may also lead to increased faith in

¹⁴² 387 U.S. 1 (1967).

¹⁴³ Cost-benefit analyses are increasingly being applied in the criminal justice realm. For example, Abrams (2013) conducts cost-benefit analyses on three types of sentencing and release policies, Abrams and Rohlfs (2011) determine optimal monetary bail amounts for felony defendants in Philadelphia using a cost-benefit analysis, Yang (2017) analyzes bail systems more generally using a cost-benefit framework, and Brown (2004) argues for a greater use of cost-benefit analyses in the criminal realm generally.

the juvenile justice system by society at large.¹⁴⁴ Such an effect could in turn encourage reporting of juvenile crimes, which may increase community safety and reduce crime in the long run by rehabilitating youth—the intended purpose of the system.¹⁴⁵

The results also indicate that a state-wide ban on indiscriminate juvenile shackling can further the rehabilitative goals of the juvenile system by increasing the probability that a juvenile is sentenced to treatment and decreasing the probability that a juvenile is sentenced to detention. This shift in case outcomes not only represents a benefit in terms of recommitting to the goals of the juvenile justice system, but it is also likely to lead to large and concrete benefits for society. Of course, the rehabilitative benefits of treatment as compared to detention will depend on the type of treatment and the detention system in the state, but research suggests that effective treatment programs can significantly reduce juvenile recidivism, whereas institutional placement may not.¹⁴⁶ Moreover, prior research indicates that effective treatment can have long-term positive effects on juveniles, such as increasing graduation rates and reducing mortality rates, which may create positive externalities for communities with at-risk youth (Heller et al. 2017). Finally, regardless of whether a decrease in the effect of appearance bias or a change in juvenile

¹⁴⁴ Previous research has illustrated that public trust and confidence in the legal system is strongly influenced by judgements about the fairness of procedures (Tyler & Fagan 2008), even more than assessments of effectiveness or beliefs about the fairness of the distribution of services (Tyler 2005).

¹⁴⁵ Studies have linked legal cynicism in communities to decreased criminal reporting and increased violent self-help alternatives, such as homicide (Kirk & Matsuda 2011; Kirk & Papachristos 2011).

¹⁴⁶ One study using longitudinal data from a large sample of serious juvenile offenders in two large cities found a null effect of institutional placement on future rates of re-arrest and self-reported offending when compared to probation (Loughran et al. 2009), while a study of juvenile correctional facilities in Ohio found that facilities that served higher risk youth, targeted dynamic risk factors with cognitive behavioral modalities, and employed trained and qualified staff were most effective at reducing recidivism (Lowenkamp et al. 2010). Conversely, meta-analyses of treatment programs typically report positive effects on recidivism (see Lipsey 1999). Specifically, some studies have found positive effects of youth summer employment programs, which decreased violence among disadvantaged youth in Chicago by 43 percent over 16 months (Heller 2014) and programs that help youth slow down and reflect on their automatic thought processes and resulting behaviors, which decreased readmission rates among juvenile detainees by 21 percent (Heller et al. 2017).

behavior explain the results in this Chapter, either mechanism represents a direct benefit from the state-wide ban on indiscriminate juvenile shackling.

Of course, any policy, no matter how beneficial, comes with some costs. The potential costs of a state-wide ban on indiscriminate juvenile shackling include (1) the cost of holding additional hearings to determine which juveniles should be shackled, (2) increased administrative and security costs to maintain order in the courtroom, (3) the costs of type I (shackling non-dangerous juveniles) and type II errors (failing to shackle dangerous juveniles), and (4) the cost of decreased deterrence due to a reduction in the seriousness of case outcomes. The cost of additional hearings will depend on the default option adopted in the statute or regulation (or by the court in the absence of clear guidance). To minimize this cost, juveniles could be presumptively unshackled absent a motion by the prosecution to hold a hearing on the issue. Alternatively, a hearing could be held in every case to determine whether shackling is appropriate. The latter method may result in fewer type II errors, which are discussed in more detail below. Given existing qualitative evidence that type II errors are relatively uncommon in states that have adopted a ban, however, the additional costs of holding a proceeding for every juvenile are probably not justified (Shaddox 2016).

The precise size of the administrative and security costs associated with bans on indiscriminate juvenile shackling may also vary between states, but they arguably do not exceed the benefits of a ban. Juveniles who are considered to be serious security threats can be shackled even after indiscriminate juvenile shackling is banned, limiting the need to hire additional security and administrative personnel. Nevertheless, a state may elect to hire more security and administrative personnel to minimize the impact of type II errors (failures to shackle dangerous defendants, discussed below). Again, these costs are likely to be minimal. According to Shaddox

(2016), New Orleans Parish, Louisiana actually reduced staffing after indiscriminate juvenile shackling was banned due to budget cuts. Despite holding roughly 4,000 juvenile hearing a year, it had had no security incidences as of 2016 (Shaddox 2016). Conversely, Clayton County in Georgia (which encompasses a portion of the greater Atlanta area) hired an additional deputy to be stationed outside the court after indiscriminate juvenile shackling was banned (Shaddox 2016). After more than a year, the deputy's services had never been required due to a lack of incidences (Shaddox 2016).

The largest costs associated with state-wide bans on indiscriminate juvenile shackling are likely to be those associated with type I and type II errors. Wrongfully ordering the shackling of a juvenile (a type I error) may be more costly than indiscriminately shackling a juvenile for a couple of reasons. First, the appearance bias effects and juvenile behavior effects of shackling may be exaggerated in this instance. Rather than appearing in shackles along with every other juvenile who is detained pretrial, the juvenile will be one of only a select few juveniles in shackles. Additionally, the juvenile will know that the court ordered them to appear in shackles specifically, which may cause them even more shame and humiliation. The shackling determination may also have an anchoring effect on future decisions in the case. If the juvenile is shackled based on a finding by the judge that he or she is dangerous, for example, a prosecutor may be less willing to negotiate a favorable plea bargain.

Failing to shackle a dangerous defendant (a type II error) can also be costly. The juvenile may escape, cause property damage, or harm another individual. Qualitative evidence from states that have passed a ban indicate that costly errors of this sort are infrequent (Shaddox 2016). For example, Miami-Dade County in Florida (which contains the city of Miami) banned indiscriminate juvenile shackling in 2006 (Shaddox 2016). As of 2016, more than 25,000

juveniles had appeared in court post-ban without injury or escape (Shaddox 2016). Similarly, Connecticut limited juvenile shackling in 2015. Post-ban, 94 percent of juveniles appeared in court without shackles, and only one escape attempt had occurred as of 2016—a youth walked out of court and turned himself in later that day (Shaddox 2016).

Nevertheless, there is a natural tradeoff between type I and type II errors that cannot be avoided. Policymakers can control this tradeoff by manipulating the amount of evidence required to justify shackling. For example, North Carolina’s ban allows for shackling in cases where a judge finds it is “reasonably necessary to maintain order, prevent the juvenile’s escape, or provide for the safety of the courtroom.”¹⁴⁷ Alternatives include a broader exception such as “whenever the judge finds that shackling is reasonable,” and a narrower exception requiring a compelling state justification, as the Court did in *Deck*.¹⁴⁸ The evidence suggesting that type II errors are infrequent indicates that a narrower exception is more appropriate. A narrower exception will also minimize type I errors and maximize the benefits that stem from a ban.

Finally, there is a possibility that a state-wide ban on indiscriminate juvenile shackling could increase juvenile crime rates by decreasing deterrence. In order for this to be the case, juveniles must be making rational decisions about whether to commit crime and be informed about the probabilities of receiving various case outcomes and how they change after the passage of a ban. There is some evidence that juvenile offenders are rational decisionmakers—in the sense that salient risks and costs of crime reduce offense rates—but mental health and developmental maturity moderate these effects (Fagan & Piquero 2007; Levitt 1998; Mocan & Rees 2005). Yet it is unclear whether changes in the distribution of sentence types, like those that result from the passage of a ban on indiscriminate juvenile shackling, would be salient enough to

¹⁴⁷ *Id.*

¹⁴⁸ 544 U.S. 622 (2005).

change behavior. Even if bans did reduce deterrence, the rehabilitative effects of increased reliance on treatment discussed above arguably cancel out at least some of the effect on crime rates.

Overall, a first pass at a cost-benefit analysis of a juvenile shackling ban indicates that the increased fairness, rehabilitation, and improved case outcomes for juveniles through either a reduction in appearance bias or changes in juvenile behavior outweigh the administrative, error-related, and deterrence costs. Even if indiscriminate juvenile shackling is not unconstitutional, states should stop focusing on whether or not a state-wide ban is necessary and start focusing on the best way to design and implement a ban to maximize social welfare, given the costs and benefits discussed above.

APPENDIX

I. Appendix A: First Survey Outline

First, all respondents viewed a consent screen and confirmed that they were over the age of 18. Then, all respondents viewed a random version of an employment hiring scenario unrelated to this dissertation. Next, all respondents viewed a random version of the vandalism trial question. There were 24 versions: the two different case variations below accompanied by a picture of the defendant who is either white or black; shackled or not; and in a suit, jeans and a t-shirt, or a jumpsuit.

A. Vandalism Variation 1

You have been selected to serve on a jury in a criminal trial. The defendant is Michael Jones, pictured below as he appears in court on the day of trial. The sole charge against Mr. Jones is vandalism, and he has pleaded not guilty. After reading the trial materials, you will be asked to decide whether the prosecution has proven that Mr. Jones is guilty of vandalism beyond a reasonable doubt.

[picture of defendant]

The object of the vandalism was a cement statue that is owned by the city and depicts the city's founder. The city had not given anyone permission to paint the statue.

The first prosecution witness was working at a store across the street from the statue. He testifies that around 9:30pm he heard someone laughing loudly. When he looked through the store window, he saw a man painting yellow pants on the statue. The witness then went outside and saw the man run over to a green truck with a broken right taillight. The man was approximately 5' 10", normal weight, and wearing a red jacket, jeans, and sneakers. The man entered the truck and sped away.

The second prosecution witness is a police officer. He testifies that he canvassed the area after receiving the above description from the eyewitness. At a house about five blocks away, he noticed a green truck with a broken right taillight, knocked on the door, and saw Mr. Jones wearing a red jacket, jeans, and sneakers. There were two other men at the house. Mr. Jones told the officer he had been with them at the house playing video games all night. Later, the police officer searched the house and found an open can of yellow paint in the basement. The police officer testifies that, based on a visual comparison, the paint in the basement is the same shade as that used on the statue. However, during cross-examination, he admits that he did not see any paint on Mr. Jones's hands or clothing.

The defense presents two alibi witnesses. They both testify that they went over to Mr. Jones's house around 8pm on the night of the incident. They made a frozen pizza and played video games with Mr. Jones for about two hours before the police officer arrived.

During closing arguments, the prosecutor stresses that Mr. Jones's clothes and car match the description given by the eyewitness, that Mr. Jones was found only five blocks from the crime scene, and that Mr. Jones had an open can of yellow paint in his basement that matches the paint used on the statue. The defense attorney stresses that the eyewitness never saw the suspect's face, that there are two individuals who verify Mr. Jones's alibi, and that there was no paint on Mr. Jones's hands or clothing.

You must apply the following rules of law:

1. Every person charged with a crime is presumed innocent.
2. Every defendant has a constitutional right not to testify. The decision not to testify cannot be used as an inference of guilt and cannot prejudice the defendant.
3. You may have to decide what testimony to believe. You may believe all of the testimony of a witness, or part of it, or none of it.
4. The prosecution has the burden of proving every element of the crime charged beyond a reasonable doubt.
5. The elements of the crime of vandalism are:
 - a. That the defendant,
 - b. knowingly,
 - c. damaged property of another,
 - d. without the owner's consent.

Which verdict do you return?

- Guilty
- Not guilty

[Page break]

How confident are you that this is the legally appropriate verdict?

- Very confident
- Somewhat confident
- Neither confident nor unconfident
- Somewhat unconfident
- Very unconfident

[Page break]

How believable did you find the following aspects of the trial?

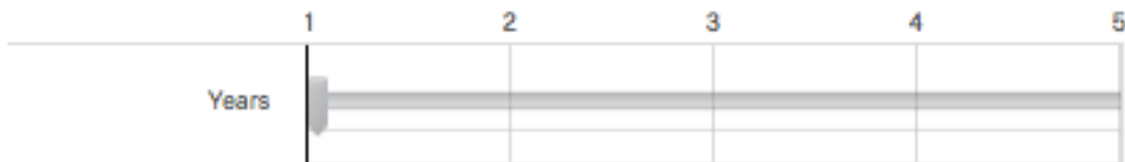
	Very believable	Somewhat believable	Neither believable nor unbelievable	Somewhat unbelievable	Very unbelievable
The prosecution's version of events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The defense's version of events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the eyewitness	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the police officer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the defendant's friends	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Page break]

Now assume that you are the judge presiding over this case. Regardless of your answers to the previous questions, assume that the jury found the defendant **guilty**.

The defendant's criminal history consists of one misdemeanor theft conviction when he was a juvenile.

You are asked to impose a sentence within the range of one to five years in prison. Your sentence can include fractions of years. What sentence do you impose?



[Page break]

How confident are you that this is the legally appropriate sentence?

- Very confident
- Somewhat confident
- Neither confident nor unconfident
- Somewhat unconfident

- Very unconfident

[Page break]

Regardless of your answers to the previous questions, assume the defendant is found ***not guilty***. How likely do you think it is that the defendant will commit a crime in the next five years?

- Very likely
- Somewhat likely
- Neither likely nor unlikely
- Somewhat unlikely
- Very unlikely

B. Vandalism Variation 2

You have been selected to serve on a jury in a criminal trial. The defendant is James Williams, pictured below as he appears in court on the day of trial. The sole charge against Mr. Williams is vandalism, and he has pleaded not guilty. After reading the trial materials, you will be asked to decide whether the prosecution has proven that Mr. Williams is guilty of vandalism beyond a reasonable doubt.

[picture of defendant]

The object of the vandalism was a community center pool. The owner had not given anyone permission to pour paint in the pool.

The first prosecution witness is a police officer. He testifies that on the night of the incident, it appears someone poured a large amount of red paint in the pool. Based on the amount of paint, he does not believe that this was an accident. The police officer also testifies that Mr. Williams consented to a search of his home, and five empty, one-gallon cans of red paint were found in his basement.

The second prosecution witness is an insurance agent. He testifies that although Mr. Williams is not the owner of the pool, he owns the aquatic equipment company that rents the pool's filtration system to the community center. Mr. Williams carries a \$50,000 insurance policy on this filtration system.

The first defense witness is Mr. Williams's wife. She testifies that he was at home with her all night on the night of the incident. She says that they made dinner and watched a movie. She also testifies that the empty paint cans are from an old home improvement project, but she can't remember why they kept the cans.

Mr. Williams also testifies on his own behalf. He admits that he owns an insurance policy on the

pool's filtration system but denies dumping the paint into the pool. He says that the empty paint cans are from an old home improvement project, but he also can't remember why they kept the cans. On the night of the incident, he says he was home with his wife, and they made dinner and watched a movie. On cross-examination, Mr. Williams admits that he currently has about \$30,000 of outstanding credit card debt that he is struggling to pay off.

During closing arguments, the prosecutor stresses the empty paint cans found in Mr. Williams's basement, the insurance policy, and Mr. Williams's financial situation. The defense attorney stresses that there is no direct evidence linking Mr. Williams to the vandalism, that the paint cans are explained by an old home improvement project, and that Mr. Williams's wife verifies his alibi.

You must apply the following rules of law:

1. Every person charged with a crime is presumed innocent.
2. You may have to decide what testimony to believe. You may believe all of the testimony of a witness, or part of it, or none of it.
3. The prosecution has the burden of proving every element of the crime charged beyond a reasonable doubt.
4. The elements of the crime of vandalism are:
 - a. That the defendant,
 - b. knowingly,
 - c. damaged property of another,
 - d. without the owner's consent.

Which verdict do you return?

- Guilty
- Not guilty

[Page break]

How confident are you that this is the legally appropriate verdict?

- Very confident
- Somewhat confident
- Neither confident nor unconfident
- Somewhat unconfident
- Very unconfident

[Page break]

How believable did you find the following aspects of the trial?

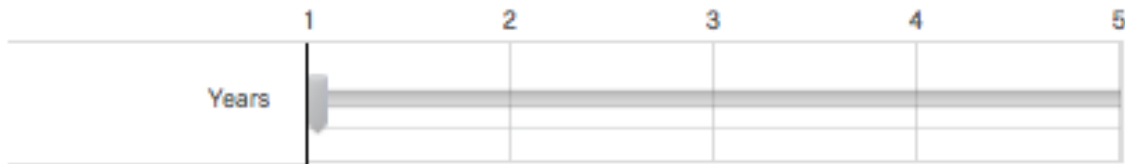
	Very believable	Somewhat believable	Neither believable nor unbelievable	Somewhat unbelievable	Very unbelievable
The prosecution's version of events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The defense's version of events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the police officer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the insurance agent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the defendant's wife	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the defendant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Page break]

Now assume that you are the judge presiding over this case. Regardless of your answers to the previous questions, assume that the jury found the defendant **guilty**.

The defendant's criminal history consists of one misdemeanor theft conviction when he was a juvenile.

You are asked to impose a sentence within the range of one to five years in prison. Your sentence can include fractions of years. What sentence do you impose?



[Page break]

How confident are you that this is the legally appropriate sentence?

- Very confident
- Somewhat confident
- Neither confident nor unconfident

- Somewhat unconfident
- Very unconfident

[Page break]

Regardless of your answers to the previous questions, assume the defendant is found ***not guilty***. How likely do you think it is that the defendant will commit a crime in the next five years?

- Very likely
- Somewhat likely
- Neither likely nor unlikely
- Somewhat unlikely
- Very unlikely

After viewing a random version of the vandalism trial question, all respondents viewed a random version of another employment hiring scenario unrelated to this dissertation. Then, all respondents viewed a version of the arson trial question. There were 24 versions: the two difference case variations below accompanied by a picture of the defendant who was either white or black; shackled or not; and in a suit, jeans and a t-shirt, or a jumpsuit. If the participant viewed the first variation of the vandalism question, they viewed a random version of the second variation of the arson question, and if the participant viewed the second variation of the vandalism question, they viewed a random version of the first variation of the arson question.

C. Arson Variation 1

You have been selected to serve on a jury in a criminal trial. The defendant is David Jackson, pictured below as he appears in court on the day of trial. The sole charge against Mr. Jackson is arson, and he has pleaded not guilty. After reading the trial materials, you will be asked to decide whether the prosecution has proven that Mr. Jackson is guilty of arson beyond a reasonable doubt.

[picture of defendant]

The fire took place inside an unoccupied building at a construction site. The property owner had not given anyone permission to start a fire.

The first prosecution witness was working at a store across the street from the construction site. He testifies that around 9:30pm he heard popping noises that sounded like fireworks. When he looked through the store window, he saw a man setting off fireworks at the construction site, which caused one of the buildings to catch fire. The witness then went outside and saw the man run towards a green truck with a broken right taillight. The man, who was laughing loudly, was approximately 5' 10", normal weight, and was wearing a red jacket, jeans, and sneakers. He entered the truck and sped away.

The second prosecution witness is a police officer. He testifies that he canvassed the area after receiving the above description from the eyewitness. At a house about five blocks away, he noticed a green truck with a broken right taillight, knocked on the door, and saw Mr. Jackson wearing a red jacket, jeans, and sneakers. There were two other men at the house. Mr. Jackson told the officer he had been with them at the house playing video games all night. Later, the police officer searched the house and found an open package of fireworks in the basement. The police officer testifies that based on his examination of the label on the fireworks, they could have caused the fire in the building. However, on cross-examination, he admits that he did not see any burn marks or soot residue on Mr. Jackson's hands or clothing.

The defense presents two alibi witnesses. They both testify that they went over to Mr. Jackson's house around 8pm on the night of the fire. They made a frozen pizza and played video games with Mr. Jackson for about two hours before the police officer arrived.

During closing arguments, the prosecutor stresses that Mr. Jackson's clothes and car match the description given by the eyewitness, that Mr. Jackson was found only five blocks from the crime scene, and that Mr. Jackson had fireworks in his basement that could have caused this fire. The defense attorney stresses that the eyewitness never saw the suspect's face, that there are two individuals who verify Mr. Jackson's alibi, and that there were no burn marks or soot residue on Mr. Jackson's hands or clothing.

You must apply the following rules of law:

1. Every person charged with a crime is presumed innocent.
2. Every defendant has a constitutional right not to testify. The decision not to testify cannot be used as an inference of guilt and cannot prejudice the defendant.
3. You may have to decide what testimony to believe. You may believe all of the testimony of a witness, or part of it, or none of it.
4. The prosecution has the burden of proving every element of the crime charged beyond a reasonable doubt.
5. The elements of the crime of arson are:
 - a. That the defendant,
 - b. By means of fire or explosives,
 - c. knowingly
 - d. damaged property of another,
 - e. without the owner's consent.

Which verdict do you return?

- Guilty
- Not guilty

[Page break]

How confident are you that this is the legally appropriate verdict?

- Very confident
- Somewhat confident
- Neither confident nor unconfident
- Somewhat unconfident
- Very unconfident

[Page break]

How believable did you find the following aspects of the trial?

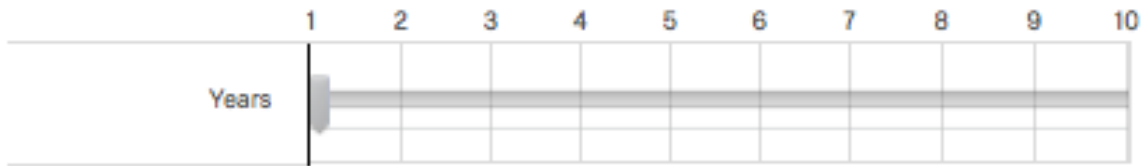
	Very believable	Somewhat believable	Neither believable nor unbelievable	Somewhat unbelievable	Very unbelievable
The prosecution's version of events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The defense's version of events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the eyewitness	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the police officer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the defendant's friends	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Page break]

Now assume that you are the judge presiding over this case. Regardless of your answers to the previous questions, assume that the jury found the defendant **guilty**.

The defendant's criminal history consists of one misdemeanor theft conviction when he was a juvenile.

You are asked to impose a sentence within the range of one to ten years in prison. Your sentence can include fractions of years. What sentence do you impose?



[Page break]

How confident are you that this is the legally appropriate sentence?

- Very confident
- Somewhat confident
- Neither confident nor unconfident
- Somewhat unconfident
- Very unconfident

[Page break]

Regardless of your answers to the previous questions, assume the defendant is found **not guilty**. How likely do you think it is that the defendant will commit a crime in the next five years?

- Very likely
- Somewhat likely
- Neither likely nor unlikely
- Somewhat unlikely
- Very unlikely

D. Arson Variation 2

You have been selected to serve on a jury in a criminal trial. The defendant is Joshua Thomas, pictured below as he appears in court on the day of trial. The sole charge against Mr. Thomas is arson, and he has pleaded not guilty. After reading the trial materials, you will be asked to decide whether the prosecution has proven that Mr. Thomas is guilty of arson beyond a reasonable doubt.

[picture of defendant]

The fire took place inside an unoccupied clothing shop. The property owner had not given anyone permission to start a fire.

The first prosecution witness is a police officer. He testifies that on the night of the incident, it

appears someone intentionally started a fire in the building using an open flame and some kind of liquid accelerant. Based on an examination of the electrical and mechanical components of the building, he does not believe the fire started accidentally. The police officer also testifies that Mr. Thomas consented to a search of his home, and five empty, 8-ounce cans of lighter fluid were found in his basement.

The second prosecution witness is an insurance agent. He testifies that although Mr. Thomas is not the owner of the building, he rents the space for his clothing shop. Mr. Thomas carries a \$50,000 insurance policy on the shop's contents.

The first defense witness is Mr. Thomas's wife. She testifies that he was at home with her all evening on the night of the fire. She says that they made dinner and watched a movie. She also testifies that the lighter fluid cans are from old camping trips, but she can't remember why they kept the cans.

Mr. Thomas also testifies on his own behalf. He admits that he owns an insurance policy on the shop's contents but denies starting the fire. He says that the lighter fluid cans are from old camping trips, but he also can't remember why they kept the cans. He testifies that he was home with his wife all evening on the night of the fire, and they made dinner and watched a movie. On cross-examination, Mr. Thomas admits that he currently has about \$30,000 of outstanding credit card debt that he is struggling to pay off.

During closing arguments, the prosecutor stresses the empty lighter fluid cans found in Mr. Thomas's basement, the insurance policy, and Mr. Thomas's financial situation. The defense attorney stresses that there is no direct evidence linking Mr. Thomas to the fire, that the lighter fluid cans are explained by old camping trips, and that Mr. Thomas's wife verifies his alibi.

You must apply the following rules of law:

1. Every person charged with a crime is presumed innocent.
2. You may have to decide what testimony to believe. You may believe all of the testimony of a witness, or part of it, or none of it.
3. The prosecution has the burden of proving every element of the crime charged beyond a reasonable doubt.
4. The elements of the crime of arson are:
 - a. That the defendant,
 - b. By means of fire or explosives,
 - c. knowingly
 - d. damaged property of another,
 - e. without the owner's consent.

Which verdict do you return?

- Guilty
- Not guilty

[Page break]

How confident are you that this is the legally appropriate verdict?

- Very confident
- Somewhat confident
- Neither confident nor unconfident
- Somewhat unconfident
- Very unconfident

[Page break]

How believable did you find the following aspects of the trial?

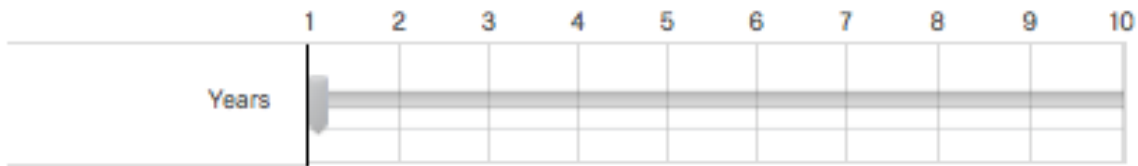
	Very believable	Somewhat believable	Neither believable nor unbelievable	Somewhat unbelievable	Very unbelievable
The prosecution's version of events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The defense's version of events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the police officer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the insurance agent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the defendant's wife	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The testimony of the defendant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Page break]

Now assume that you are the judge presiding over this case. Regardless of your answers to the previous questions, assume that the jury found the defendant **guilty**.

The defendant's criminal history consists of one misdemeanor theft conviction when he was a juvenile.

You are asked to impose a sentence within the range of one to ten years in prison. Your sentence can include fractions of years. What sentence do you impose?



[Page break]

How confident are you that this is the legally appropriate sentence?

- Very confident
- Somewhat confident
- Neither confident nor unconfident
- Somewhat unconfident
- Very unconfident

[Page break]

Regardless of your answers to the previous questions, assume the defendant is found **not guilty**. How likely do you think it is that the defendant will commit a crime in the next five years?

- Very likely
- Somewhat likely
- Neither likely nor unlikely
- Somewhat unlikely
- Very unlikely

Finally, all respondents answered a series of demographic questions, including questions about gender, age, race, ethnicity, family, residence, education, income, political preference, dedication to religion, prior arrest experience, prior criminal conviction experience, prior jury service, prior crime victimization, height, weight, eating and exercise habits, and prior employee hiring experience. There were also two opinion questions that asked participants to rate how well they think the criminal justice system works in the United States and how big of a problem they think employment discrimination still is in the United States.

II. Appendix B: Second Survey Outline

First, all respondents viewed a consent screen and confirmed that they were over the age of 18. Then, all respondents viewed a random version of the criminal scenario question. There were 48 different versions: a crime of either vandalism or arson with a picture of one of four

defendants, two of whom were black and two of whom were white, shackled or not, and in a suit, jeans and a t-shirt, or a jumpsuit.

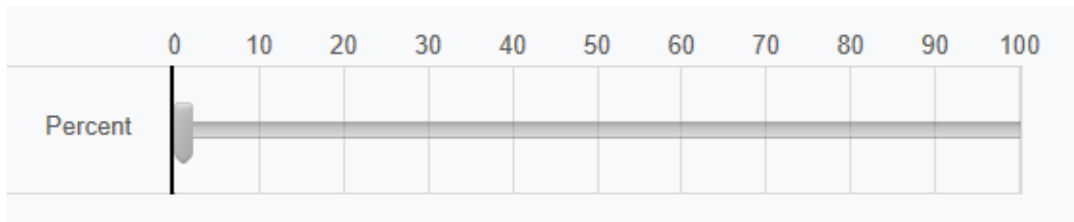
A. Crime Question Version 1

You have been selected to serve on a jury in a criminal trial. The defendant is [**Michael Jones/James Williams**], pictured below as he appears in court on the day of trial. The sole charge against Mr. [**Jones/Williams**] is vandalism, and he has pleaded not guilty. At the end of the trial, you will be asked to apply the following rules of law:

1. Every person charged with a crime is presumed innocent.
2. You may decide what testimony to believe. You may believe all of the testimony of a witness, or part of it, or none of it.
3. The prosecution has the burden of proving every element of the crime beyond a reasonable doubt.
4. The elements of the crime of vandalism are:
 1. That the defendant,
 2. knowingly,
 3. damaged property of another,
 4. without the owner's consent.

[*picture of defendant*]

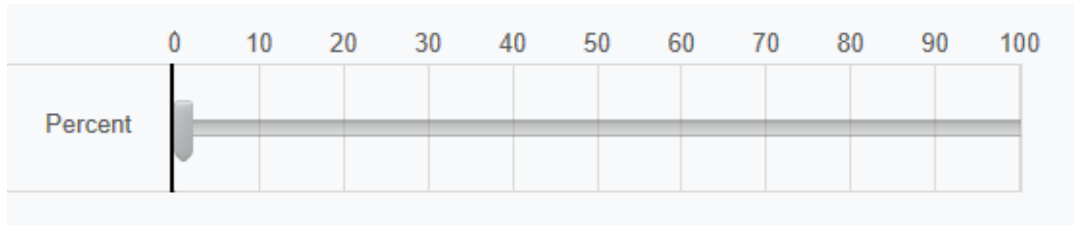
Before hearing any evidence, how likely do you think it is that the defendant is guilty?



Please rate the importance of the factors below in answering the previous question.

	Very Important	Important	Neutral	Little Importance	Not at All Important
The Crime Charged	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Defendant's Propensity to Commit Crime	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Your Opinions About the Criminal Justice System	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

After hearing the evidence, how sure will you need to be that the defendant is guilty in order to return a verdict of guilty?



Please rate how important the factors below will be in deciding whether to convict.

	Very Important	Important	Neutral	Little Importance	Not at All Important
Fear of a Wrongful Conviction	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fear of a Wrongful Acquittal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Whether the Prosecution Has Met the Burden of Proof	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

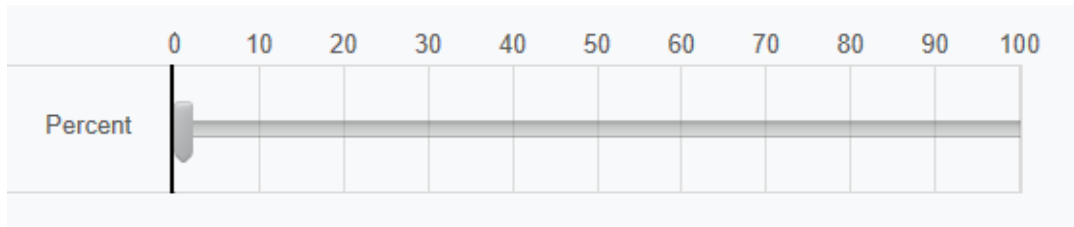
B. Crime Question Version 2

You have been selected to serve on a jury in a criminal trial. The defendant is [**David Jackson/Joshua Thomas**], pictured below as he appears in court on the day of trial. The sole charge against Mr. [**Jackson/Thomas**] is arson, and he has pleaded not guilty. At the end of the trial, you will be asked to apply the following rules of law:

1. Every person charged with a crime is presumed innocent.
2. You may decide what testimony to believe. You may believe all of the testimony of a witness, or part of it, or none of it.
3. The prosecution has the burden of proving every element of the crime beyond a reasonable doubt.
4. The elements of the crime of arson are:
 1. That the defendant,
 2. by means of fire or explosives,
 3. knowingly,
 4. damaged property of another,
 5. without the owner's consent.

[*picture of defendant*]

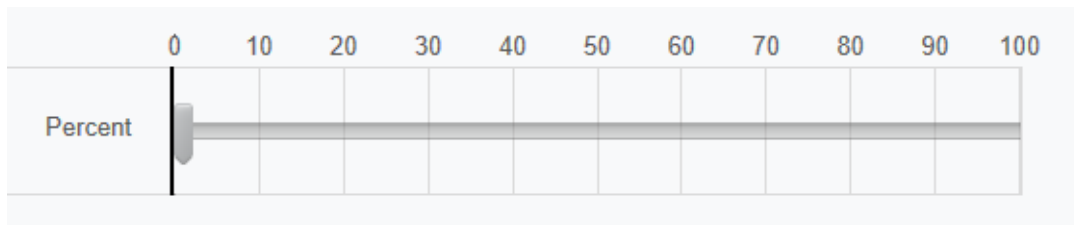
Before hearing any evidence, how likely do you think it is that the defendant is guilty?



Please rate the importance of the factors below in answering the previous question.

	Very Important	Important	Neutral	Little Importance	Not at All Important
The Crime Charged	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Defendant's Propensity to Commit Crime	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Your Opinions About the Criminal Justice System	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

After hearing the evidence, how sure will you need to be that the defendant is guilty in order to return a verdict of guilty?



Please rate how important the factors below will be in deciding whether to convict.

	Very Important	Important	Neutral	Little Importance	Not at All Important
Fear of a Wrongful Conviction	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fear of a Wrongful Acquittal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Whether the Prosecution Has Met the Burden of Proof	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Next, all the respondents saw a random version of the attractiveness question. There were 24 different versions: a picture of one of four defendants, two who were black and two who were white, shackled or not, and in a suit, jeans and a t-shirt, or a jumpsuit. Participants never saw a picture of the same defendant that they saw in the crime question.

C. Attractiveness Question

[*defendant picture*]

How would you rate the attractiveness of the individual above?

- Very attractive
- Somewhat attractive
- Neither attractive nor unattractive
- Somewhat unattractive
- Very unattractive

Finally, all respondents answered a series of demographic questions, including questions about gender, age, race, ethnicity, residence, education, income, political preference, dedication to religion, prior arrest experience, prior criminal conviction experience, prior jury service, and prior crime victimization. There was also an opinion question that asked participants to rate how well they think the criminal justice system works in the United States.

APPENDIX TABLES

Table A1. First Study Subject Demographics by Race/Ethnicity

Characteristic	White, Non-Hispanic	Hispanic	Black	Asian
Female	56.6%	45.4%	63.5%	41.9%
Age	36.3	32.5	34.8	32.3
US Native	97.8%	90.5%	93.7%	69.2%
Married	44.2%	38.6%	28.4%	34.4%
B.A. or Higher (If 25+)	57.2%	52.3%	51.6%	82.0%
Employed	82.3%	85.0%	82.1%	81.0%
Median Household Income (\$2017)	\$52,500	\$52,500	\$37,500	\$52,500
Urban	28.9%	49.1%	48.7%	47.6%
Northeast	20.1%	16.2%	16.6%	18.2%
Midwest	24.4%	9.6%	15.6%	11.9%
South	37.1%	42%	55.0%	27.7%
West	18.4%	31.4%	12.6%	42.3%
Democrat	39.9%	47.4%	63.2%	57.0%
Republican	26.1%	17.6%	5.9%	14.0%
Religious	21.7%	23.8%	38.5%	28.1%
Prior Jury Service	13.6%	13.1%	16.6%	11.5%
Criminal History	17.8%	19.7%	16.2%	6.1%
Criminal Victimization	43.0%	39.5%	39.1%	22.8%
Distrusts Justice System	42.3%	45.6%	57.9%	35.6%

Notes: Age is calculated for employed persons only; values are calculated for individuals who report only a single race. Individuals who attend church at least once a month are considered religious. Individuals who have a prior criminal conviction or have been arrested are considered to have a criminal history. Individuals are considered to distrust the justice system if they strongly disagree or somewhat disagree with the statement that the criminal justice system works well in the United States.

Table A2. Second Study Subject Demographics by Race/Ethnicity

Characteristic	White, Non-Hispanic	Hispanic	Black	Asian
Female	54.9%	51.0%	59.9%	50.3%
Age	36.4	32.3	33.7	34.1
US Native	97.8%	92.9%	92.5%	61.7%
Married	48.0%	47.6%	33.5%	50.9%
B.A. or Higher (If 25+)	59.7%	65.0%	51.4%	77.2%
Employed	84.5%	89.1%	83.5%	81.4%
Median Household Income (\$2017)	\$52,500	\$52,500	\$37,500	\$67,500
Urban	33.5%	61.2%	54.2%	42.5%
Northeast	17.8%	12.6%	14.6%	21.6%
Midwest	23.4%	6.5%	13.7%	8.4%
South	41.6%	44.9%	56.1%	28.1%
West	17.1%	33.7%	15.6%	41.9%
Democrat	38.6%	38.5%	58.7%	43.4%
Republican	34.1%	30.1%	9.7%	13.8%

Religious	32.8%	42.7%	41.7%	31.1%
Prior Jury Service	16.7%	20.4%	17.0%	15.6%
Criminal History	18.9%	25.8%	17.5%	7.2%
Criminal Victimization	44.3%	38.0%	36.7%	26.9%
Distrusts Justice System	31.0%	28.2%	47.2%	25.7%

Notes: Age is calculated for employed persons only; values are calculated for individuals who report only a single race. Individuals who attend church at least once a month are considered religious. Individuals who have a prior criminal conviction or have been arrested are considered to have a criminal history. Individuals are considered to distrust the justice system if they strongly disagree or somewhat disagree with the statement that the criminal justice system works well in the United States.

Table A3. Case Demographics for Cases Involving Juveniles Ages 6–15

Variable	Tennessee	North Carolina
Number of Cases	264,967	226,040
Age at Referral	13.48	13.72
Age at Offense	—	13.65
Female	36.07%	29.26%
White, Non-Hispanic	56.60%	38.85%
White, Hispanic	1.30%	6.07%
Black, Non-Hispanic	39.85%	50.89%
Asian, Non-Hispanic	0.32%	0.52%
Native American, Non-Hispanic	0.08%	1.72%
Multi-Racial or Non-White, Hispanic	1.85%	1.95%
Living with Biological Parents	16.01%	—
Living with Remarried Parent	6.51%	—
Living with Single Parent	57.80%	—
Living with Relatives	8.47%	—
Living with Adopted Parents	0.71%	—
Placed Living Situation	3.47%	—
Living Independently	0.31%	—
Other Living Situation	1.16%	—
Enrolled in School	97.08 %	—
Enrolled in Special Education Program	9.41%	—
Weapon Used	—	15.95%
Firearm/ Explosive Used	—	1.77%
Serious Injury	—	1.95%
School-Related Case	—	33.05%
Pretrial Detention	19.43%	—
Felony Charged	—	14.28%
Misdemeanor Charged	—	75.88%
Number of Initial Charges	1.20	1.94
Multiple Initial Charges	15.39%	72.45%
Charged with Crime Against Person	17.18%	36.00%
Charged with Property Crime	19.98%	33.01%
Charged with Drug or Alcohol Crime	5.30%	8.37%
Charged with Public Order Crime	27.63%	17.55%
Charged with Status Offense	37.38%	14.89%

Length of Case (In Months)	2.63	1.95
Dismissed	37.91%	55.13%
Transferred to Adult Court	0.06%	0.17%
Sentenced	52.47%	43.64%
Missing Case Outcome	9.55%	1.06%
Harshest Sentence - Detention	7.29%	23.63%
Harshest Sentence - Probation	39.86%	60.27%
Harshest Sentence - Treatment	18.49%	12.47%
Harshest Sentence - Other	34.36%	1.73%
Missing Sentence	—	1.90%

Notes: Observations are dropped from the Tennessee data if they do not involve any criminal charge, and observations are dropped from both states if they only involve traffic offenses or minor infractions.

Observations from the North Carolina data are assigned the mean age at offense if the juvenile is over 17 at the time of the offense or over 15 at the time of the offense and is not charged with at least one status offense, and observations from the North Carolina data are recoded so that the age at offense is equal to the age at referral if the age at offense is greater than the age at referral. Case length is treated as missing if it is negative. Missing values for non-case outcome variables that are available in both states, pretrial detention, and felony charged are replaced with regression-estimated values, and missing values for variables only available in one state (other than pretrial detention and felony charged) are not included in this table. The sentencing percentages are calculated only for the sample of cases that result in a sentence. The charge categories are not mutually exclusive, except in the nationwide data, which only consider the most serious charge.

REFERENCES

- Abrams, David S. 2013. "The Prisoner's Dilemma: A Cost-Benefit Approach to Incarceration." *Iowa Law Review*, 98: 905–969.
- Abrams, David S., Marianne Bertrand, and Sendhil Mullainathan 2012. "Do Judges Vary in Their Treatment of Race?" *The Journal of Legal Studies*, 41(2): 347–384.
- Abrams, David S. and Chris Rohlfs 2011. "Optimal Bail and the Value of Freedom: Evidence from the Philadelphia Bail Experiment." *Economic Inquiry*, 49: 750–770.
- Abwender, David A. and Kenyatta Hough 2001. "Interactive Effects of Characteristics of Defendant and Mock Juror on U.S. Participants' Judgment and Sentencing Recommendations." *The Journal of Social Psychology*, 141(5): 603–615.
- Adams, Kenneth and Charles R. Cutshall 1987. "Refusing to Prosecute Minor Offenses: The Relative Influence of Legal and Extralegal Factors." *Justice Quarterly*, 4(4): 595–609.
- Alesina, Alberto and Eliana La Ferrara 2014. "A Test of Racial Bias in Capital Sentencing." *American Economic Review*, 104(11): 3397–3433.
- Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson 2012. "The Impact of Jury Race in Criminal Trials." *The Quarterly Journal of Economics*, 127(2): 1017–1055.
- Baker, Al, J. David Goodman, and Benjamin Mueller 2015. "Beyond the Chokehold: The Path to Eric Garner's Death." *New York Times*, <https://www.nytimes.com/2015/06/14/nyregion/eric-garner-police-chokehold-staten-island.html>.
- Baldus, David C., George G. Woodworth, and Charles A. Pulaski Jr. 1990. *Equal Justice and the Death Penalty: A Legal and Empirical Analysis*. Northeastern University Press.
- Baldwin, John and Michael McConville 1980. "Juries, Foremen and Verdicts." *British Journal of Criminology*, 20(1): 35–44.
- Balganesh, Shyamkrishna, Irina D. Manta, and Tess Wilkinson-Ryan 2014. "Judging Similarity." *Iowa Law Review*, 100: 267–290.
- Bassett, Debra L. 2006. "The Rural Venue." *Alabama Law Review*, 57: 941–974.
- Baumer, Eric P., Steven F. Messner, and Richard B. Felson 2000. "The Role of Victim Characteristics in the Disposition of Murder Cases." *Justice Quarterly*, 17(2): 281–307.
- Becker, Gary S. 1957. *The Economics of Discrimination*. 2nd Edition, University of Chicago Press.
- Berggren, Niclas, Henrik Jordahl, and Panu Poutvarra 2017. "The Right Look: Conservative

- Politicians Look Better and Voters Reward It.” *Journal of Public Economics*, 146: 79–86.
- Bernard, J. L. 1979. “Interaction Between the Race of the Defendant and That of Jurors in Determining Verdicts.” *Law and Psychology Review*, 5: 103–111.
- Beyer, Marty 2015. *Affidavit*, <http://njdc.info/wp-content/uploads/2014/09/Beyer-Affidavit-w-CV-Jan-2015-Final.pdf>.
- Biddle, Jeff E. and Daniel S. Hamermesh 1998. “Beauty, Productivity, and Discrimination: Lawyers’ Looks and Lucre.” *Journal of Labor Economics*, 16(1): 172–201.
- Bidwell, Robert 2015. *Affidavit*, <http://njdc.info/wp-content/uploads/2014/09/Bidwell-Shackling-Affidavit-General-April-2015.pdf>.
- Binnall, James M. 2017. “Summoning Criminal Desistance: Convicted Felons’ Perspectives on Jury Service.” *Law and Social Inquiry*, 43(1): 4–27.
- Birkhead, Tamar R. 2013. “Closing the Widening Net: The Rights of Juveniles at Intake.” *Texas Tech Law Review*, 46: 157–186.
- Bischoff, Bea 2018. “When Being a Good Lawyer Means Dressing Your Clients.” *Racked*, <https://www.racked.com/2018/1/18/16900864/public-defender-clothing-accused-donations>.
- Blumstein, Alfred 1993. “Racial Disproportionality of US Prison Populations Revisited.” *University of Colorado Law Review*, 64: 743–760.
- Bottoms, Bette L., Suzanne L. Davis, and Michelle A. Epstein 2004. “Effects of Victim and Defendant Race on Jurors’ Decisions in Child Sexual Abuse Cases.” *Journal of Applied Social Psychology*, 34(1): 1–33.
- Bridges, Virginia 2018. “In Durham, Children Are Often Restrained in Court. Is it a Violation of State Law?” *Herald Sun*, <https://www.heraldsun.com/news/local/article220505815.html>.
- Broker, Phyllis J. and Kevin J. Reddington 2014. “The Defendant’s Case.” In *Massachusetts Superior Court Criminal Practice Manual*, ch. 14.
- Brown, Darryl K. 2004. “Cost-Benefit Analysis in Criminal Law.” *California Law Review*, 92(2): 323–372.
- Brown, Michael J. 2008. “The Effects of Eyeglasses and Race on Juror Decisions Involving a Violent Crime.” *American Journal of Forensic Psychology*, 26(2): 25–43.
- Bureau of Justice Statistics 2017. “Federal Justice Statistics, 2014 – Statistical Tables,” <https://www.bjs.gov/content/pub/pdf/fjs14st.pdf>.
- Bureau of Justice Statistics 2013. “Felony Defendants in Large Urban Counties, 2009 –

- Statistical Tables,” <https://www.bjs.gov/content/pub/pdf/fdluc09.pdf>.
- Bureau of Labor Statistics 2018. “Occupational Outlook Handbook: Judges and Hearing Officers,” <https://www.bls.gov/ooh/legal/judges-and-hearing-officers.htm>.
- Butler, Paul D. 1997. “Race-Based Jury Nullification: Case-in-Chief.” *John Marshall Law Review*, 30: 911–922.
- Card, David and Alan B. Krueger 1994. “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania.” *American Economics Review*, 84: 772–793.
- Cheng, Edward K. 2013. “Reconceptualizing the Burden of Proof.” *The Yale Law Journal*, 122: 1254–1279.
- Cohn, Ellen S., et al. 2009. “Reducing White Juror Bias: The Role of Race Salience and Racial Attitudes.” *Journal of Applied Social Psychology*, 39(8): 1953–1973.
- Conley, John M., William J. Turnier, and Mary R. Rose 2000. “The Racial Ecology of the Courtroom: An Experimental Study of Juror Response to the Race of Criminal Defendants.” *Wisconsin Law Review*, 1185–1220.
- Culhane, Scott E., Harmon M. Hosch, and William G. Weaver 2004. “Crime Victims Serving as Jurors: Is There Bias Present?” *Law and Human Behavior*, 28(6): 649–659.
- Cryder, Cynthia, Simona Botti, and Yvetta Simonyan 2017. “The Charity Beauty Premium: Satisfying Donors’ “Want” Versus “Should” Desires.” *Journal of Marketing Research*, LIV: 605–618.
- Davis, James A., et al. 1975. “The Decision Processes of 6- and 12-Person Mock Juries Assigned Unanimous and Two-Thirds Majority Rules.” *Journal of Personality and Social Psychology*, 32(1): 1–14.
- DeSantis, Andrea and Wesley A. Kayson 1997. “Defendants’ Characteristics of Attractiveness, Race, and Sex and Sentencing Decisions.” *Psychological Reports*, 8: 679–683.
- Dixon, Travis L. and Keith B. Maddox 2005. “Skin Tone, Crime News, and Social Reality Judgments: Priming the Stereotype of the Dark and Dangerous Black Criminal.” *Journal of Applied Social Psychology*, 35(8): 1555–1570.
- Dumas, Rafaele and Benoit Teste 2006. “The Influence of Criminal Facial Stereotypes on Juridic Judgements.” *Swiss Journal of Psychology*, 65: 237–244.
- Efran, Michael G. 1974. “The Effect of Physical Appearance on the Judgment of Guilt, Interpersonal Attraction, and Severity of Recommended Punishment in a Simulated Jury Task.” *Journal of Research in Personality*, 8: 45–54.
- Elek, Jennifer K. and Paula Hannaford-Agor 2014. “Can Explicit Instructions Reduce

- Expressions of Implicit Bias? New Questions Following a Test of a Specialized Jury Instruction.” *National Center for State Courts*.
- Enough, Birte and Thomas Mussweiler 2001. “Sentencing Under Uncertainty: Anchoring Effects in the Courtroom.” *Journal of Applied Social Psychology*, 31(7): 1535–1551.
- Fagan, Jeffrey and Alex R. Piquero 2007. “Rational Choice and Developmental Influences on Recidivism Among Adolescent Felony Offenders.” *Journal of Empirical Legal Studies*, 4(4):715–746.
- Federal Bureau of Investigation 2017. “Crime in the United States,” <https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017>.
- Feld, Barry C. 2014. *Juvenile Justice Administration in a Nutshell*. 3rd Edition, West Economic Publishing.
- Field, Hubert S. 1979. “Rape Trials and Jurors’ Decisions: A Psycholegal Analysis of the Effects of Victim, Defendant, and Case Characteristics.” *Law and Human Behavior*, 3(4): 261–284.
- Fisher, Talia 2012. “Conviction Without Conviction.” *Minnesota Law Review*, 96: 833–855.
- Fletcher, Jason M. 2009. “Beauty vs. Brains: Early Labor Market Outcomes of High School Graduates.” *Economics Letters*, 105: 321–325.
- Foley, Linda A. and Minor H. Chamblin 1982. “The Effect of Race and Personality on Mock Jurors’ Decisions.” *The Journal of Psychology*, 112: 47–51.
- Fontaine, Gary and Rick Kiger 1978. “The Effects of Defendant Dress and Supervision on Judgments of Simulated Jurors: An Exploratory Study.” *Law and Human Behavior*, 2: 63–71.
- Ford, Julian D. 2014. *Affidavit*, <http://njdc.info/wp-content/uploads/2014/09/Ford-Affidavit-Final-Dec-2014.pdf>.
- Frank, Hannah 2017. “Unambiguous Deterrence: Ambiguity Attitudes in the Juvenile Justice System and the Case for a Right to Counsel During Intake Proceedings.” *Vanderbilt Law Review*, 70(2):709–735.
- Free, Marvin D. Jr. 2001. “Racial Bias and the American Criminal Justice System: Race and Presentencing Revisited.” *Critical Criminology*, 10(3): 195–223.
- Freitas, Sarah 1995. “Extending the Privilege Against Self-Incrimination to the Juvenile Waiver Hearing.” *University of Chicago Law Review*, 62: 301–329.
- Friend, Ronald M. and Michael Vinson 1974. “Leaning Over Backwards: Jurors’ Responses to Defendants’ Attractiveness.” *Journal of Communication*, 24(3): 124–129.

- Frieze, Irene H., Josephine E. Olson, and June Russell 1991. "Attractiveness and Income for Men and Women in Management." *Journal of Applied Social Psychology*, 21(13): 1039–1057.
- Funk, Friederike and Alexander Todorov 2013. "Criminal Stereotypes in the Courtroom: Facial Tattoos Affect Guilt and Punishment Differently." *Psychology, Public Policy, and Law*, 19(4): 466–478.
- Galarza, Francisco B. and Gustavo Yamada 2016. "Triple Penalty in Employment Access: The Role of Beauty, Race, and Sex." *Journal of Applied Economics*, XX(1): 29–47.
- Gallaher, Brian D. and John C. Lore III 2008. "Shackling Children in Juvenile Court: The Growing Debate, Recent Trends, and the Way to Protect Everyone's Interest." *University of California Davis Juvenile Justice Law and Policy*, 12: 453–480.
- Ginther, Matthew R., et al. 2014. "The Language of Mens Rea." *Vanderbilt Law Review*, 67: 1327–1372.
- Gordon, Randall A. 1993. "The Effect of Strong Versus Weak Evidence on the Assessment of Race Stereotypic and Race Nonstereotypic Crimes." *Journal of Applied Social Psychology*, 23(9): 734–749.
- Gray, David B. and Richard D. Ashmore 1976. "Biasing Influence of Defendants' Characteristics on Simulated Sentencing." *Psychological Reports*, 38: 727–738.
- Greenwald, Anthony G. and Linda H. Krieger 2006. "Implicit Bias: Scientific Foundations." *California Law Review*, 94(4): 945–967.
- Griffin, Eugene 2014. *Affidavit*, <http://njdc.info/wp-content/uploads/2014/09/Griffin-Affidavit-II.pdf>.
- Gupta, Nabanita D., Nancy L. Etcoff, and Mads M. Jaeger 2016. "Beauty in Mind: The Effects of Physical Attractiveness on Psychological Well-Being and Distress." *Journal of Happiness Studies*, 17: 1313–1325.
- Guthrie, Chris, Jeffrey Rachlinski, and Andrew J. Wistrich 2007. "Blinking on the Bench: How Judges Decide Cases." *Cornell Law Review*, 93: 1–43.
- Hamermesh, Daniel S. 2006. "Changing Looks and Changing "Discrimination": The Beauty of Economists." *Economics Letters*, 93(3): 405–412.
- Hamermesh, Daniel S. and Jason Abrevaya 2013. "Beauty Is the Promise of Happiness?" *European Economic Review*, 64: 351–368.
- Hamermesh, Daniel S. and Jeff E. Biddle 1994. "Beauty and the Labor Market." *The American Economic Review*, 84(5): 1174–1194.
- Hamermesh, Daniel S., Xin Meng, and Junsen Zhang 2002. "Dress for Success—Does Primping

- Pay?" *Labour Economics*, 9: 361–373.
- Hamermesh, Daniel S. and Amy Parker 2005. "Beauty in the Classroom: Instructors' Pulchritude and Putative Pedagogical Productivity." *Economics of Education Review*, 24: 369–376.
- Hannaford-Agor, et al. 2003. "Evaluation of Hung Juries in Bronx County, New York, Los Angeles County, California, Maricopa County, Arizona, and Washington, DC, 2000-2001." Ann Arbor, MI: Inter-university Consortium for Political and Social Research, <https://doi.org/10.3886/ICPSR03689.v1>.
- Harrigan, Jinni A., et al. 1996. "Detecting Anxiety and Defensiveness from Visual and Auditory Cues." *Journal of Personality*, 64(3): 657–709.
- Hatfield, Elaine and Susan Sprecher 1986. *Mirror, Mirror . . . : The Importance of Looks in Everyday Life*. SUNY Press.
- Heller, Sara B. 2014. "Summer Jobs Reduce Violence Among Disadvantaged Youth." *Science*, 346(6214): 1219–1223.
- Heller, Sara B., et al. 2017. "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago." *Quarterly Journal of Economics*, 132(1): 1–54.
- Hernandez-Julian, Rey and Christina Peters 2017. "Student Appearance and Academic Performance." *Journal of Human Capital*, 11(2): 247–262.
- Hersch, Joni and Beverly Moran 2015. "Coitus and Consequences in the Legal System: An Experimental Study." *Southern Methodist University Law Review*, 68: 927–949.
- Hersch, Joni and Jennifer B. Shinall 2017. "Something to Talk About: Information Exchange Under Employment Law." *University of Pennsylvania Law Review*, 165: 49–90.
- Hill, Erick L. and Jeffrey E. Pfeifer 1992. "Nullification Instructions and Juror Guilt Ratings: An Examination of Modern Racism." *Contemporary Social Psychology*, 16: 6–10.
- Hockenberry, Sarah and Charles Puzzanchera 2018. "Juvenile Court Statistics 2018." *National Center for Juvenile Justice*, <https://www.ojjdp.gov/ojstatbb/njcda/pdf/jcs2016.pdf>.
- Hurwitz, John and Mark Peffley 1997. "Public Perceptions of Race and Crime: The Role of Racial Stereotypes." *American Journal of Political Science*, 41(2): 375–401.
- Hymes, Robert W., et al. 1993. "Acquaintance Rape: The Effect of Race of Defendant and Race of Victim on White Juror Decisions." *The Journal of Social Psychology*, 133(5): 627–634.
- Ingriselli, Elizabeth 2015. "Mitigating Jurors' Racial Biases: The Effects of Content and Timing of Jury Instructions." *Yale Law Journal*, 124: 1690–1740.

- Jacobson, Steven K. and Charles R. Berger 1974. "Communication and Justice: Defendant Attributes and Their Effects on the Severity of his Sentence." *Communication and Justice*, 41(3): 282–286.
- Jones, Todd R. and Joseph Price 2017. "Information and the Beauty Premium in Political Elections." *Contemporary Economic Policy*, 35(4): 677–683.
- Kalt, Brian C. 2003. "The Exclusion of Felons from Jury Service." *American University Law Review*, 53: 65–189.
- Kerwin, Jeffrey and David R. Shaffer 1994. "Mock Jurors Versus Mock Juries: The Role of Deliberations in Reactions to Inadmissible Testimony." *Personality and Social Psychology Bulletin*, 20(2): 153–162.
- Kirk, David S. and Mauri Matsuda 2011. "Legal Cynicism, Collective Efficacy, and the Ecology of Arrest." *Criminology*, 49(2): 443–472.
- Kirk, David S. and Andrew V. Papachristos 2011. "Cultural Mechanisms and the Persistence of Neighborhood Violence." *American Journal of Sociology*, 116(4): 1190–1233.
- Koosed, Margery M. 2001. "Averting Mistaken Executions by Adopting the Model Penal Code's Exclusion of Death in the Presence of Lingering Doubt." *Northern Illinois University Law Review*, 21: 41–129.
- Kramer, Geoffrey P., Norbert L. Kerr, and John Carroll 1990. "Pretrial Publicity, Judicial Remedies, and Jury Bias." *Law and Human Behavior*, 14(5): 409–438.
- Kraus, Louis J. 2015. *Affidavit*, <http://njdc.info/wp-content/uploads/2014/09/Kraus-Affidavit-General-Final.pdf>.
- Kuziemko, Ilyana, et al. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *American Economic Review*, 105(4): 1478–1508.
- Lassiter, Daniel G. 2002. "Illusory Causation in the Courtroom." *Current Directions in Psychological Science*, 11(6): 204–208.
- Lee, Jean N. 2017. "The Process is the Punishment: Juror Demographics and Case Administration in State Courts." *American Law and Economics Review*, 19(2): 361–390.
- Leigh, Andrew and Tirta Susilo 2008. "Is Voting Skin-Deep? Estimating the Effect of Candidate Ballot Photographs on Election Outcomes." *Journal of Economic Psychology*, 30: 61–70.
- Leipold, Andrew D. 1996. "The Dangers of Race-Based Jury Nullification: A Response to Professor Butler." *UCLA Law Review*, 44: 109–141.
- Leventhal, Gloria and Ronald Krate 1977. "Physical Attractiveness and Severity of Sentencing." *Psychological Reports*, 40: 315–318.

- Levinson, Justin D., Huajian Cai, and Danielle Young 2010. "Guilty By Implicit Racial Bias: The Guilty/Not Guilty Implicit Association Test." *Ohio State Journal of Criminal Law*, 8: 187–208.
- Levitt, Steven D. 1998. "Juvenile Crime and Punishment." *Journal of Political Economy*, 106: 1156–1185.
- Lipsey, Mark W. 1999. "Can Rehabilitative Programs Reduce the Recidivism of Juvenile Offenders – An Inquiry into the Effectiveness of Practical Programs." *Virginia Journal of Social Policy and the Law*, 6: 611–642.
- London, Kamala and Narina Nunez 2000. "The Effect of Jury Deliberation on Jurors' Propensity to Disregard Inadmissible Evidence." *Journal of Applied Psychology*, 85(6): 932–939.
- Loughran, Thomas A. 2009. "Estimating a Dose-Response Relationship Between Length of Stay and Future Recidivism in Serious Juvenile Offenders." *Criminology*, 47(3): 699–740.
- Lowenkamp, Christopher T., et al. 2010. "Community Corrections Facilities for Juvenile Offenders in Ohio: An Examination of Treatment Integrity and Recidivism." *Criminal Justice and Behavior*, 37(6): 695–708.
- MacCoun, Robert J. 1990. "The Emergence of Extra-Legal Bias During Jury Deliberation." *Criminal Justice and Behavior*, 17(3): 303–314.
- Mazzella, Ronald and Alan Feingold 1994. "The Effects of Physical Attractiveness, Race, Socioeconomic Status, and Gender of Defendants and Victims on Judgments of Mock Jurors: A Meta-Analysis." *Journal of Applied Social Psychology*, 24(15): 1315–1338.
- McGuire, Mary V. and Gordon Bermant 1977. "Individual and Group Decisions in Response to a Mock Trial: A Methodological Note." *Journal of Applied Social Psychology*, 7(3): 220–226.
- McLaurin, Kim M. 2012. "Children in Chains: Indiscriminate Shackling of Juveniles." *Washington University Journal of Law and Policy*, 38: 213–239.
- Miller, Monica K., et al. 2011. "The Effects of Deliberations and Religious Identity on Mock Jurors' Verdicts." *Group Processes & Intergroup Relations*, 14(4): 517–532.
- Mills, Carol J. and Wayne E. Bohannon 1980. "Juror Characteristics: To What Extent Are They Related to Jury Verdicts?" *Judicature*, 64(1): 22–31.
- Mitchell, Tara L., et al. 2005. "Racial Bias in Mock Juror Decision-Making: A Meta-Analytic Review of Defendant Treatment." *Law and Human Behavior*, 29(6): 621–637.
- Mobius, Markus M. and Tanya S. Rosenblat 2006. "Why Beauty Matters." *American Economic Review*, 96(1): 222–235.
- Mocan, Nanci H. and Daniel I. Rees 2005. "Economic Conditions, Deterrence and Juvenile

- Crime: Evidence from Micro Data.” *American Law and Economics Review*, 7: 319–349.
- Mocan, Nanci H. and Erdal Tekin 2010. “Ugly Criminals.” *The Review of Economics and Statistics*, 92(1): 15–30.
- Moore, Martha T. 2007. “Should Kids Go to Court in Chains?” *USA Today*, http://www.usatoday.com/news/nation/2007-06-17-shackles_N.htm.
- Myers, Martha A. 1982. “Common Law in Action: The Prosecution of Felonies and Misdemeanors.” *Sociological Inquiry*, 52(1): 1–15.
- Nabha, Anita 2008. “Shuffling to Justice: Why Children Should Not Be Shackled in Court.” *Brooklyn Law Review*, 73: 1549–1589.
- Nakell, Barry and Kenneth A. Hardy 1987. *The Arbitrariness of the Death Penalty*. Temple University Press.
- National Juvenile Defender Center. “Issue Brief: Campaign Against Indiscriminate Juvenile Shackling,” http://njdc.info/wp-content/uploads/2016/01/NJDC_CAIJS_Issue-Brief.pdf.
- National Registry of Exonerations, <https://www.law.umich.edu/special/exoneration/Pages/about.aspx>.
- Nellis, Ashley 2016. “The Color of Justice: Racial and Ethnic Disparity in State Prisons.” *The Sentencing Project*.
- Nemeth, Charlan and Ruth Hyland Sosis 1973. “A Simulated Jury Study: Characteristics of the Defendant and the Jurors.” *The Journal of Social Psychology*, 90(2): 221–229.
- Ostrom, Thomas M., et al. 1978. “An Integration Theory Analysis of Jurors’ Presumptions of Guilt or Innocence.” *Journal of Personality and Psychology*, 36(4): 436–450.
- Page, Antony 2005. “Batson’s Blind-Spot: Unconscious Stereotyping and the Peremptory Challenge” *Boston University Law Review*, 85: 155–262.
- Patry, Marc W. 2008. “Attractive but Guilty: Deliberation and the Physical Attractiveness Bias.” *Psychological Reports*, 102: 727–733.
- Peffley, Mark, Todd Shields, and Bruce Williams 1996. “The Intersection of Race and Crime in Television News Stories: An Experimental Study.” *Political Communication*, 13(3): 309–327.
- Pennington, Nancy and Reid Hastie 1991. “A Cognitive Theory of Juror Decision Making: The Story Model.” *Cardozo Law Review*, 13: 519–558.
- Pennington, Nancy and Reid Hastie 1981. “Juror Decision-Making Models: The Generalization Gap.” *Psychological Bulletin*, 89(2): 246–287.

- Pfeifer, Jeffrey E. and James R. P. Ogloff 1991. "Ambiguity and Guilt Determinations: A Modern Racism Perspective." *Journal of Applied Social Psychology*, 21: 1713–1725.
- Piehl, Jochen 1977. "Integration of Information in the "Courts:" Influence of Physical Attractiveness on Amount of Punishment for a Traffic Offender." *Psychological Reports*, 41: 551–556.
- Pope, Carl E. and William Feyerherm 1995. "Research Summary: Minorities and the Juvenile Justice System." *U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention*.
- Price, Michael K. 2008. "Fund-Raising Success and Solicitor's Beauty Capital: Do Blondes Raise More Funds?" *Economics Letters*, 100: 351–354.
- Puzzanchera, Charles, Benjamin Adams, and Melissa Sickmund 2011. "Juvenile Court Statistics 2008." *National Center for Juvenile Justice*, <http://www.ncjj.org/pdf/jcsreports/jcs2008.pdf>.
- Rachlinski, Jeffrey J. and Sheri L. Johnson 2009. "Does Unconscious Racial Bias Affect Trial Judges?" *Notre Dame Law Review*, 84(3): 1195–1246.
- Rachlinski, Jeffrey J., Andrew J. Wistrich, and Chris Guthrie 2015. "Can Judges Make Reliable Numeric Judgements? Distorted Damages and Skewed Sentences." *Indiana Law Journal*, 90: 695–739.
- Rosenblitt, Donald 2015. *Affidavit*, <http://njdc.info/wp-content/uploads/2014/09/Rosenblitt-Affidavit-Notarized-CV-Final-1-6-15.pdf>.
- Rossner, Meredith, et al. 2017. "Dock on Trial: Courtroom Design and the Presumption of Innocence." *Journal of Law and Society*, 44(3): 317–344.
- Rule, Nicholas O. and Nalini Ambady 2008. "The Face of Success: Inferences from Chief Executive Officers' Appearance Predict Company Profits." *Psychological Science*, 19(2): 109–111.
- Russell, Joyce 2017. "Iowa Chief Justice Bans Routine Shackling of Kids in Court." *Iowa Public Radio*, <http://iowapublicradio.org/post/iowa-chief-justice-bans-routine-shackling-kids-court#stream/0>.
- Ruva, Christine, Cathy McEvoy, and Judith B. Bryant 2007. "Effects of Pre-Trial Publicity and Jury Deliberation on Juror Bias and Source Memory Errors." *Applied Cognitive Psychology*, 21(1): 45–67.
- Salyers, Lance 1999. "Invaluable Tool vs. Unfair Use of Private Information: Examining Prosecutors' Use of Jurors' Criminal History Records in Voir Dire." *Washington and Lee Law Review*, 56: 1079–1124.

- Schvey, Natasha A., et al. 2013. "The Influence of a Defendant's Body Weight on Perceptions of Guilt." *International Journal of Obesity*, 37: 1275–1281.
- Sentencing Project 2008. "Reducing Racial Disparity in the Criminal Justice System: A Manual for Practitioners and Policymakers."
- Shaddox, Colleen 2016. "Shackling and Courtroom Safety." *Campaign Against Indiscriminate Juvenile Shackling*, https://njdc.info/wp-content/uploads/2016/03/CAIJS_Shackling-and-Courtroom-Safety-3.4.16.pdf.
- Sickmund, Melissa 2005. "Delinquency Cases in Juvenile Court, 2005." *Office of Juvenile Justice and Delinquency Prevention*, <https://www.ncjrs.gov/pdffiles1/ojjdp/224538.pdf>.
- Sigall, Harold and Nancy Ostrove 1975. "Beautiful but Dangerous: Effects of Offender Attractiveness and Nature of Crime on Juridic Judgment." *Journal of Personality and Social Psychology*, 31(3): 410–414.
- Skolnick, Paul and Jerry I. Shaw 1997. "The OJ Simpson Criminal Trial Verdict: Racism or Status Shield?" *Journal of Social Issues*, 53(3): 503–516.
- Smith, Edward D. and Anita Hed 1979. "Effects of Offenders' Age and Attractiveness on Sentencing by Mock Juries." *Psychological Reports*, 44(3): 691–694.
- Sommers, Samuel R. 2006. "Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations." *Journal of Personality and Social Psychology*, 90: 597–612.
- Sommers, Samuel R. and Pheobe Ellsworth 2000. "Race in the Courtroom: Perceptions of Guilt and Dispositional Attitudes." *Personality and Social Psychology Bulletin*, 26(11): 1367–1379.
- Sorenson, Todd A., et al. 2014. "Do You Receive a Lighter Prison Sentence Because You Are a Woman or a White? An Economic Analysis of the Federal Criminal Sentencing Guidelines." *B.E. Journal of Economic Analysis and Policy*, 14(1): 1–54.
- Spohn, Cassia 2000. "Thirty Years of Sentencing Reform: The Quest for a Racially Neutral Sentencing Process." *Policies, Processes, and Decisions of the Criminal Justice System*, 3: 427–501.
- Sunnafrank, Michael and Norman E. Fontes 1983. "General and Crime Related Racial Stereotypes and Influence on Juridic Decision." *The Cornell Journal of Social Relations*, 17: 1–15.
- Swaine, Jon 2015. "Walter Scott Shooting: Officer Laughs About Adrenaline Rush in Recording." *The Guardian*, <https://www.theguardian.com/us-news/2015/apr/12/walter-scott-shooting-officer-michael-slager-audio-recording>.
- Tyler, Tom R. 2005. "Policing in Black and White: Ethnic Group Differences in Trust and

- Confidence in the Police.” *Police Quarterly*, 8(3): 322–342.
- Tyler, Tom R. and Jeffrey Fagan 2008. “Legitimacy and Cooperation: Why Do People Help the Police Fight Crime in Their Communities.” *Ohio State Journal of Criminal Law*, 6: 231–276.
- U.S. Census Bureau a. “Genealogy: Frequently Occurring Surnames from the Census 2000,” https://www.census.gov/topics/population/genealogy/data/2000_surnames.html.
- U.S. Census Bureau b. “Quick Facts: North Carolina,” <https://www.census.gov/quickfacts/fact/table/nc/PST045217>.
- U.S. Census Bureau c. “Quick Facts: Tennessee,” <https://www.census.gov/quickfacts/fact/table/tn/EDU635216>.
- U.S. Census Bureau d. “Quick Facts: United States,” <https://www.census.gov/quickfacts/fact/table/tn/EDU635216>.
- Weitzer, Ronald and Steven A. Tuch 2005. “Determinants of Public Satisfaction with the Police.” *Police Quarterly*, 8(3): 279–297.
- Wilkinson-Ryan, Tess and Jonathan Baron 2008. “The Effect of Conflicting Moral and Legal Rules on Bargaining Behavior: The Case of No-Fault Divorce.” *The Journal of Legal Studies*, 37(1): 315–338.
- Wistrich, Andrew J., Jeffrey J. Rachlinski, and Chris Guthrie 2015. “Heart Versus Head: Do Judges Follow the Law or Follow Their Feelings?” *Texas Law Review*, 93: 855–923.
- Wurm, Gwen 2015. *Affidavit*, <http://njdc.info/wp-content/uploads/2015/01/Gwen-Wurm-full-shackling-affidavit-Jan-2015.pdf>.
- Yang, Crystal S. 2017. “Toward an Optimal Bail System.” *New York University Law Review*, 92: 1399–1493.