EMPIRICAL ESSAYS ON THE ECONOMICS OF BURDENS OF PROOF

by

Ye Yuan

Copyright © Ye Yuan 2021

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ECONOMICS

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2021
THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by: Ye Yuan
titled: Empirical Essays on the Economics of Burdens of Proof

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Tiemen Woutersen  
Date: 8/6/2021

Katherine Barnes  
Date: 8/12/2021

Price Fishback  
Date: 8/13/2021

Juan Pantano  

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Tiemen Woutersen  
Date: 8/6/2021

Dissertation Committee Chair  
Economics Department
Acknowledgment

I am deeply grateful for the mentorship and guidance from my advisors: Tiemen Woutersen, Kathie Barnes, Price Fishback, and Juan Pantano. I hope you will find this dissertation worthy of your time and efforts in training me to become an economist. I coauthored Chapter 1 of this dissertation with Price Fishback and Chapter 2 with Spencer Cooper. I thank both for being wonderful research collaborators. You set a very high bar for productive research collaboration. For helpful discussions at various stages of my research, I thank Ashley Langer, Evan Taylor, Jessamyn Schaller, Gautam Gowrisankaran, Marc Miller, Dan Klerman, JJ Prescott, Alison Morantz, Sarah Gotschall and LuAnn Haley. I also thank the seminar participants at the University of Arizona Department of Economics, the American Law and Economics Association Annual Meeting, and the QuantLaw Conference at the University of Arizona College of Law. Your skepticism and thoughtful comments greatly improved my work.

I had never thought about becoming an empirical researcher until taking Kathie’s criminal procedure class in 2013 as an 1L. As a law student, I was drawn to the empirical analysis of law because it can bring important clarity to legal discussions that typically lead to “no right answers.” Five years’ academic training in economics helped me develop the quantitative skills to study questions that interest me. I also learned how to make valid empirical claims based on evidence rather than imagination. Hopefully, this dissertation and my future work can demonstrate that I am able to combine my academic training in law and empirical methods to produce decent research.

I would never have accomplished this without the support and encouragement from my parents, who inspired me to dream big, work hard, and be a smart problem solver. This dissertation is dedicated to you.
# Table of Contents

**Abstract** ————————————————————————————————————————————————————11

**Chapter 1: Quantifying the Monetary Impacts of Nonquantitative Changes in Liability and Procedural Rules: A Study of Workers’ Compensation, 1997 – 2016** ————13

1. Introduction ———————————————————————————————————————————14
2. The Structure of Workers’ Compensation ————————————————————17
3. Recent State Legislation ————————————————————————————20
   3.1 Overview ————————————————————————————————————20
   3.2 Liberal Construction of Statutes ————————————————————21
   3.3 Intoxication Presumed as a Proximate Cause of Injury ————22
   3.4 Benefit Apportionment Due to Preexisting Conditions ————23
   3.5 Medical-Care Restriction Laws ———————————————————24
4. Measuring the Monetary Value of Changes in Burden of Proof in Liability Settings ————————————————————————26
5. Data ——————————————————————————————————————————28
   5.1 Cash Benefits and Medical Benefits ———————————————————29
   5.2 Workers’ Compensation Expected Benefits Index ————29
   5.3 Fatal Accident Rate —————————————————————————31
   5.4 State Laws —————————————————————————————————32
7. Results for the Burden of Proof Index ———————————————————35
8. Results for Each Type of the Laws and Decomposition Analysis ————37
9. Conclusion ————————————————————————————————————40

**Chapter 2: Racial, Gender Disparities and State Prosecutors’ Discretion: Evidence from Blakely v. Washington** ————60

1. Introduction ————————————————————————————————————61
2. Institutional Background ————————————————————64
   2.1 Sentencing in North Carolina ———————————————————64
   2.2 *Blakely v. Washington* ———————————————————66
# List of Figures

<table>
<thead>
<tr>
<th>Figure 1.1</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.2</td>
<td>43</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>44</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>45</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>46</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>47</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>48</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>49</td>
</tr>
<tr>
<td>Figure 1.9</td>
<td>50</td>
</tr>
<tr>
<td>Figure 1.10</td>
<td>51</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>92</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>93</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>94</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>95</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>96</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>97</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>98</td>
</tr>
<tr>
<td>Figure 2.8</td>
<td>99</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>118</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>153</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>154</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>155</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>156</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>157</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>158</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>159</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>160</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>161</td>
</tr>
</tbody>
</table>
List of Tables

Table 1.1 ..................................................................................................................... 52
Table 1.2 ..................................................................................................................... 53
Table 1.3 ..................................................................................................................... 54
Table 1.4 ..................................................................................................................... 55
Table 1.5 ..................................................................................................................... 56
Table 1.6 ..................................................................................................................... 57
Table 1.7 ..................................................................................................................... 58
Table 1.8 ..................................................................................................................... 59
Table 2.1 ..................................................................................................................... 100
Table 2.2 ..................................................................................................................... 101
Table 2.3 ..................................................................................................................... 102
Table 2.4 ..................................................................................................................... 103
Table 2.5 ..................................................................................................................... 104
Table 2.6 ..................................................................................................................... 105
Table 2.7 ..................................................................................................................... 106
Table 3.1 ..................................................................................................................... 162
Table 3.2 ..................................................................................................................... 163
Table 3.3 ..................................................................................................................... 164
Table 1.A.1 ............................................................................................................... 165
Table 1.A.2 ............................................................................................................... 166
Table 1.A.3 ............................................................................................................... 167
Table 1.A.4 ............................................................................................................... 168
Table 1.A.5 ............................................................................................................... 169
Table 1.A.6 ............................................................................................................... 170
Table 1.A.7 ............................................................................................................... 171
Table 1.A.8 ............................................................................................................... 172
Table 1.A.9 ............................................................................................................... 173
Table 1.A.10 ............................................................................................................ 174
Abstract

Chapter 1 examines the effects of state-level workers’ compensation reforms on workers’ access to compensation benefits. In nearly all liability settings the probability that an accident victim will receive payments and the amount of those payments are strongly influenced by laws or rules related to burden of proof, legal procedure, access to medical treatment, and other nonquantitative factors. Yet, their impact is often difficult to measure empirically because it is difficult to measure the true liability and size of damages because those are also at issue in the proceedings. We study the impact of such nonquantitative laws on average accident payments per covered worker under workers’ compensation, where fault is not at issue and damage payments are set by statute. The results show that states experienced sizeable drops in cash and medical payments per covered worker of 20 and 16 percent, respectively, after they adopted burden of proof laws. A decomposition of national means between 1997 and 2016 shows that about 11 percent of the decline in the national average of benefits per covered worker might be attributed to the patterns of adoption of these laws. The analysis also shows that average statutory expected benefits rose during this period and thus served to counteract the decline in benefits per covered worker.

Chapter 2 investigates the causal effects of restricting prosecutorial discretion on racial and gender disparities. In North Carolina, sentence enhancements may be imposed if prosecutors meet the legally required burden of showing the existence of aggravating factors. Blakely v. Washington 542 U.S. 296 (2004) imposed a significant constraint on state prosecutors’ discretion by raising their burdens of proof in successfully seeking sentence enhancements from “preponderance of evidence” to “beyond a reasonable doubt.” First, we use a regression discontinuity design to exploit the timing of the Blakely decision and find that Blakely decreases the defendants’ likelihood of receiving sentence enhancements by 66%. Second, we find striking evidence that restricting prosecutorial discretion disproportionately benefits male defendants and eliminates the entire unexplained pre-Blakely gender gap of men being 28% more likely to receive sentence enhancements than women. Examining subgroups of defendants with the same charging offenses, we find that Blakely disproportionally reduces the likelihood of receiving sentence enhancements among men charged with assault and women charged with forgery. This sharp comparison suggests that prosecutors tend to associate defendants’ gender with different types of aggravating factors and pursue sentence enhancements under weak evidence. In contrast, we find no evidence suggesting a racial gap of Black or Hispanic defendants being more likely to receive sentence enhancements than non-Black and non-Hispanic counterparts both pre and post Blakely. This finding provides no evidence suggesting prosecutors’
Chapter 3 examines, both theoretically and empirically, the effects of the Court’s enforcement of the element rule. The element rule, proposed by Stephanos Bibas, means that facts increasing defendants’ statutory maximum sentences should be proved to juries beyond a reasonable doubt. We know very little about how the element rule affects criminal prosecution. This chapter tries to fill this important gap. First, I proposed a conceptual framework on how the element rule affects the prosecutors’ behavior. The model predicts a “direct effect” which makes prosecutors less likely to pursue sentence enhancements because complying with the “beyond a reasonable doubt” standard increases the costs and decreases the expected benefits of pursuing enhancements. Enforcing the element rule may also generate spillover effects because prosecutors may compensate the loss of opportunities to punish defendants with sentence enhancements by changing other prosecution decisions. Secondly, I empirically estimated the direct and spillover effects by using large administrative criminal case records from North Carolina. I proxied the Court’s enforcement of the element rule using the Supreme Court’s decision, Blakely v. Washington. Through a regression discontinuity framework, I found strong evidence of the direct effects that Blakely reduced defendants’ likelihood of receiving sentence enhancements by 66%. On the other hand, I also found abundant evidence on the spillover effects of enforcing the element rule. Prosecutors responded to Blakely by becoming more stringent to defendants in making many prosecution decisions, such as pleading guilty from a felony to a misdemeanor. Furthermore, I found no evidence that Blakely significantly changed defendant’s minimum sentence lengths, suggesting that the spillover effects of the element rule may have offset the its direct effects in protecting defendants against illegitimate punishment. The results highlight that broad prosecutorial discretion may undermine the effectiveness of external regulation of criminal prosecution.
Chapter 1

I Introduction

The American legal system provides monetary transfers to compensate people for harms caused by a variety of torts, including workplace accidents, medical malpractice, product malfunction, transportation accidents, and breach of contractual obligations. Not only are these transfers influenced by liability rules themselves, but a variety of procedural rules determine burdens of proof and adjust for extenuating circumstances. Such rules are non-quantitative and we expect that they have impacts on the probability of compensation and the amount received. It is difficult, however, to determine the size of that impact when parties are trying to assess the expected payments (probability times amount received) when making their decisions about the case. One way to do it is to compile data sets on payments with and without the rules in place. Even after collecting the data, it is difficult to get precise estimates of the impact of the procedural rules for most liability cases because assigning liability is still at issue. Controlling for liability is difficult in many situations because it is hard to quantify the factors that determine liability. Further, there are no standardized damage awards. Consider death awards, where severity is not an issue. In many cases, opposing expert witnesses provide different calculations that are not readily available to researcher. The final decision maker can choose among these calculations or ignore them in the large number of out-of-court settlements or in the small share of cases decided by courts. Thus, it is difficult to control for two key factors that determine the size of the award when trying to isolate the impact of procedural rules.

The workers’ compensation system offers an opportunity to control both for liability and standardized damages when trying to determine the monetary consequences of changes in procedural rules. Problems associated with controlling for liability of employer or the worker are narrowed significantly because workers’ compensation is a strict liability system and the employer is generally liable for all accidents “arising out of or in the course of employment.” Further, the statutes in each state set a standardized set of benefit calculations for a variety of accident severities. Therefore, we can control for the standardized benefits when estimating the impact of the procedural rules on expected compensation from the workplace accidents.

The workers’ compensation system therefore allows us to develop better estimates of how large an effect procedural rules have on changes in expected damages. Workers’ compensation practi-
tioners, scholars and the press have expressed concerns that benefit payments for covered workers have declined in part due to a series changes in state procedural laws that influenced the probability of receiving any compensation and/or prevented workers from getting full statutory compensation [2]. A Department of Labor (undated, 17) report suggested that such changes “had enormous, though perhaps more hidden, impact on injured workers’ access to benefits.”

The view that workers’ compensation benefits have been eroding has been heavily influenced by two types of information. The first is the decline in the inflation-adjusted national averages of cash and medical benefits per covered worker shown in Figure 1.1. The second is an index of workers’ compensation laws for each year between 2002 and 2014 created by Yue Qiu and Michael Grabell in 2015. In an online article based on the index, Grabell and Berkes (2015) came to the conclusion that many states had “slashed workers’ compensation benefits,” and were in a “race to the bottom.” One plausible reason for the declines seen in Figure 1.1 is a significant drop in workplace accident rates during the period, which benefited the workers.

To understand better how the changes in workers’ compensation laws have influenced payments per covered worker, we start with the information compiled by Qiu (2015). Rather than focusing on their aggregate index, we examine the key changes in the components of their index. We then use state statutes and compendiums to identify the laws in place when we start the study in 1997 and to flesh out information on the changes in law between 1997 and 2016. Our analysis redid the portions of the Pro Publica measures of legislative changes in compensation benefits generosity. Pro Publica’s description of changes in workers’ compensation benefits is problematic because it does not adequately take into account the fact that vast majority of the states chose to index their weekly maximum benefit payments with state-level average wages or inflation. In these “indexing” states, the weekly maximums increase every year without any legal amendment and such benefit increases are not shown in Pro Publica’s measure. In Section V.2, we fix this problem by developing an expected benefits index that incorporates state-level statutory benefits for workers who suffer different types of injuries.

We analyze the impacts of five different types of laws, all of which can plausibly influence both cash and medical benefits: (1) laws that prohibit the use of the liberal construction rule; (2) laws that presume intoxication was the proximate cause of injuries to an intoxicated worker; (3) laws

---

that apportion benefits based on a workers’ preexisting conditions; (4) laws establishing medical fee schedules, and (5) laws establishing stat-wide networks of physicians to treat work-related injuries. The first three changes raise workers’ burdens of proof and inevitably require workers to seek additional evidence to support their claims, which would impose additional costs on their attempts to obtain benefits or even make obtaining benefits infeasible. The last two legal changes primarily aim at reducing medical costs. However, cash benefits might also be plausibly affected by these two changes if they lead to shorter treatment periods of work-related injuries (Bernacki et al. 2006).

To measure the impact of these laws, we develop an annual panel data set for the states between 1997 and 2016. The outcome variables in the empirical analysis are cash and medical benefit payments per covered worker. We employ a generalized difference-in-difference empirical strategy that exploits the inter-state and inter-temporal variation in the passage of new workers’ compensation laws. We also estimate event study models to test the parallel trends assumption and to show the evolutionary paths of the treatment effects over time. The empirical analysis also incorporates the recent econometric methods developed to ensure that the time-varying treatment effects do not bias the estimates from two-way fixed effects and event-study models (Borusyak and Jaravel 2017; Goodman-Bacon 2018; Sun and Abraham 2020; De Chaisemartin and d’Haultfoeuille 2020A; de Chaisemartin and D’Haultfoeuille 2020B). We apply the methodologies developed in De Chaisemartin and d’Haultfoeuille (2020A) and de Chaisemartin and D’Haultfoeuille (2020B) to estimate the dynamic treatment effects and also the weights associated with the two-way fixed effects estimates.

First, we create an index aggregating three different types of burden-of-proof laws and find that an additional burden-of-proof law is associated with a reduction in cash and medical benefit payments per covered worker by 20 and 16 percent respectively. A decomposition analysis suggests that the burden of proof laws can explain around 11 percent of the drops in the national averages of both cash and medical benefit payments per covered worker. The event study analysis further shows that the estimated effects are driven by the passage of the laws, instead of preexisting trends. We also find evidence on the dynamic effects of the laws. The negative effect of the laws on cash benefit payments is 11 percent in the first year after the passage and increases to 24 percent, more than double, in the third year. The laws’ impacts on reducing medical benefit payments do not fully manifest until the second year after the passage and are persistent afterwards.

Second, from the statutes on benefit payments, we develop present value calculations of statutory payments made for four types of accidents. Then, using fixed accident rates as weights, we
calculate an expected benefit index for each state and year. The index provides a much more precise picture of the changes in statutory benefits. Figure 1.2 shows an increasing trend in the index value, which paints a quite different picture of why average benefits per eligible worker have changed in Figure 1.1 from the one depicted by the Pro Publica measure. The main reason for the difference is that most of the states index their benefits to average wages or inflation. The estimated elasticity of cash benefit payments with respect to statutory benefits is around 1.2, which is consistent with previous findings in the literature (Krueger and Burton Jr. 1990).

Third, we examine separately the effects of each of the five workers' compensation laws on benefits payments. The findings suggest that the laws prohibiting liberal construction and laws apportioning benefits based on preexisting conditions lead to significant decrease in both cash and medical benefit payments. Physician network laws lead to a significant decrease in cash benefit payments but the negative effects on medical benefits appear to be driven by a preexisting trend. Benefits are negatively correlated with intoxication laws but the event study analysis suggests a similar preexisting trend. We do not find evidence suggesting any significant impacts of establishing medical fee schedule laws on either cash or medical benefit payments per worker covered.

The remainder of the paper is organized as follows. Section II reviews relevant institutional features and the economic literature of workers’ compensation programs. Section III provides detailed descriptions of each of the five types of laws we empirically examine. Section IV presents a conceptual framework of measuring the monetary value of changes in burden of proof law. Section V describes the data source. Section VI presents the generalized difference-in-difference econometric models. Sections VII and VIII describe the results of the effects of the burden of proof index and each types of the laws we examine. Section X offers concluding remarks.

II The Structure of Workers’ Compensation

From their beginnings workers’ compensation programs have been state programs. Employers are liable for accidents to workers that “arise out of or in the course of employment.” Employers cover this liability by direct purchase of insurance or by forming self-insured pools if they can document adequate resources. Workers’ compensation benefits generally consist of two types: medical ben-

---

3In many states employers purchase from private insurance companies. In 2019 North Dakota, Ohio, Washington, and Wyoming operate exclusive state compensation fund, although Ohio and Washington allow self-insurance in limited cases. Some private insurance can be found in those states in cases where group policies overlap state boundaries. Another 21 states have insurance funds that compete with private insurers for workers’ compensation coverage. For more see: https://www.insureon.com/insurance-glossary/workers-compensation-state-fund

17
efits and cash benefits. Medical benefits cover the entire scope of medical expenses incurred in treating work-related injuries and illnesses. Cash benefits partially compensate for lost earnings of workers related to the accident. Such benefits are available to workers who became disabled due to workplace accidents or to the dependents of deceased workers. The wage-replacement benefits typically have a duration limit and are a fraction of workers’ pre-injury wages, with maximum and minimum restrictions usually tied to state average wages. Overall, these state safety net programs distribute billions of dollars. In 2015 the total amount of workers’ compensation benefits payments to non-federal employees amounted to $30 billion for medical payment and $28 billion for wage replacement.

Economic studies of workers’ compensation focus on examining the impacts of the following three key regulations: monetary benefits, non-monetary eligibility requirements and medical-care restrictions. The main outcome variables are related to workplace safety and the program costs. The empirical strategy typically exploits the inter-state and inter-temporal variation of regulations of state workers’ compensation programs.

One large strand of literature used a variety of measures of the cash benefits for wage replacement, which we refer to as “expected benefits.” Allen (2015) and Fishback and Kantor (1995) developed measures of cash benefits to examine the political economy of the states’ choice of benefits across time and place. Studies used measures of cash benefits to examine their influence on injury rates (Butler and Worrall 1983; Moore and Viscusi 1989; Krueger 1990; Ruser 1993; Guo and Burton Jr 2010). Similarly, a few papers looked at the effects of cash benefits on program costs and the estimated “benefit elasticity” of costs with respect to cash benefits (Krueger and Burton Jr 1990; Guo and Burton Jr 2010). Overall, this strand of literature has focused on the period from the 1970s to the 1990s. One of our contributions is to update the estimate of the “benefit elasticity” using data for a more recent time frame, from 1997 to 2016.

A smaller strand of literature developed measures that capture non-monetary eligibility requirements for injured workers to receive compensation benefits. Burton and Spieler (2001) highlighted the legislative phenomenon initiated in the 1990s that tightened the eligibility requirements for workers’ compensation benefits. A few studies documented significant impacts of imposing eligibility requirements on both workplace safety and cash benefit payments (Boden and Ruser 2003; Guo and Burton Jr 2010; Gentry and Viscusi 2019). None of the studies to date have examined the structure of medical payments nor have they examined how specific procedural issues and burdens of proof might change the size of the benefits. What’s more, prior work typically focused on
eligibility laws that were passed in the 1980s and 1990s.

One of our novel contribution comes from examining how benefit payments per covered worker in the states are influenced by laws related to burden of proof and proximate causes of injuries. Raising workers’ burdens of proof is a more subtle procedural reform, instead of directly altering the amounts to be paid. Such procedural modifications can be a powerful channel affecting substantive outcomes. Congressional Representative John Dingell once said: “I'll let you write the substance...you let me write the procedure, and I’ll screw you every time.” In a comprehensive treatise on workers’ compensation, [Spieler (2016)] argues that “issues of proof” and the uses of legislative presumptions meant to streamline the administrative process have reduced worker access to benefits. The U.S. Department of Labor (undated, 17 and 22) shares the same concerns:

“Changes in the processing and adjudication of claims have had enormous, though perhaps more hidden, impact on injured workers’ access to benefits.”

“New features of workers’ compensation systems, such as higher burdens of proof for injured workers, serve to reduce access to benefits.”

We assess how large an effect these changes in laws have had on changes in benefits per covered worker, both cash and medical.

Another contribution is our updating of the studies of the impact of medical-care restriction laws. Compared with injuries that occurred in non-work environments, the medical costs of treating work-related injuries are much higher (Durbin et al. 1996; Butler and Worrall 1983). In response, states have been passing laws that restrict the use of medical care, such as imposing medical fee schedules and limiting workers’ free choice of treating physicians. Earlier studies have examined the effects of medical-care restriction laws on work-related injuries and medical benefit payments but did not find evidence suggesting that the laws were effective (Boden and Fleischman 1989; Pozzebon 1994; Boden and Ruser 2003). None of the existing studies measured the impacts of medical-care restriction laws that were passed in the past two decades. Our research also fills

4 The exception is Gentry and Viscusi (2019) who used the Pro Publica index and developed two dummies which reflect whether workers’ compensation programs become more generous or stringent. The dummy variables broadly include a wide range of eligibility rules that were passed in the recent two decades. The Pro Publica index only focuses on the changes in laws between 2002 and 2014 and does not include laws of the same type that were in place at the start of the period. We used the Pro Publica laws as a starting point and then did a comprehensive search of the workers’ compensation laws already in place at the start of our time period and then went through the statutes that followed to double check the Pro Publica laws and to add any that they may have missed.

5 Regulatory Reform Act: Hearing on H.R. 2327 Before the H. Subcomm. on Admin. Law and Governmental Relations, 98th Cong. 312 (1983)
this gap by assessing how cash and medical benefits are affected by medical fee schedule laws and physician network laws.

III Recent State Legislation

III.1 Overview

Under the workers’ compensation system, the burden of proof usually rests upon the injured worker to establish two types of causation: legal and medical. To establish legal causation, workers are usually required to make a *prima facie* showing all of the following three elements: (1) employees acted in the course of employment; (2) employees suffered an injury from an accident arising out of and in the course of employment; and (3) work activities contributed to the injury. To establish medical causation, workers are generally required to show that the accident caused the injury. Usually, the evidentiary standard required to show causation is by the “preponderance of the evidence.” If the worker fails to establish any of the three elements of causation, the benefits can be reduced or eliminated. Showing causation does not guarantee benefit payments. Employers could still avoid the payment by raising statutory defenses, such as intoxication and preexisting conditions.

The workers’ actual burdens are highly contextual and vary significantly based on the type of injury, the workers’ characteristics, and the governing state law. The states have different presumptions of compensability, which affect the allocation of burdens of proof between employers and workers. For example, Maine case law establishes a presumption that work injuries are not caused by diseases unrelated to employment. As a result, if employers attempt to escape compensation liability on the ground that the injuries are caused by prior diseases, they bear the burden to establish such a causal link. What’s more, the workers’ burdens of proof could also change due to individual-specific characteristics, like the presence of a preexisting medical condition. An injured worker with a preexisting condition might have to show that the workplace accident was the major contributing cause of his injuries.

In examining the laws identified by *Pro Publica* and through our own reading of the statutes, we focus on five types of laws that influenced the likelihood of compensation or might reduce the amount of compensation on a proportional basis: (1) prohibitions on the use of the liberal construction rule;

---


7 *Davidson v. Bancroft*, 560 A2d 13
(2) presumption of workers’ intoxication as the proximate cause of their injuries; (3) apportionment of benefits on the basis of workers’ preexisting medical conditions; (4) the establishment of medical fee schedules; and (5) laws that initialize and regulate physician network to treat work-related injuries. Table 1.1 summarizes the states and times when each type of the laws was passed.

III.2 Liberal Construction of Statutes

Awards of benefits involve the interpretation of a variety of statutes. The “liberal construction rule” interprets the statute by presuming that workers prevail by default if there are ambiguities related to the merits of the claim. The countervailing doctrine is the “strict construction rule,” which requires adjudicators to resolve disputes solely based on the evidentiary merits of the claims and opposes presuming either party as the default winner.

Statutes and case laws are the two main ways for states to adopt a liberal construction rule. Ohio and Maryland, for example, have explicit statutes that require the law to be liberally construed.\(^8\) The most commonly seen justification favoring the liberal construction rule is the remedial nature of workers compensation laws. For example, a Wisconsin decision held that a firefighter who died while saving his family from a fire at night was acting within the scope of employment; therefore, his family could receive workers’ compensation death benefits. One justification for the ruling is that the remedial purpose of the law requires the statutory term, “within scope of employment,” to be liberally interpreted to allow benefits payments.\(^9\)

On the other hand, states prohibiting the liberal construction rule argue against the remedial nature of workers’ compensation law. In those states, the goal of workers’ compensation laws is to strike a balance between workers’ interests in obtaining benefits and employers’ interests in maintaining reasonably low costs of compensating work injuries. Statutes in Florida\(^10\) and Minnesota\(^11\) explicitly say that workers’ compensation mutually renounces common law rights and defenses of workers and employers. Thus, the liberal construction rule is disallowed because it unjustifiably places workers’ interests over employers’ interests.

Prior to 1997, the beginning of our sample period, seven states prohibited the use of the liberal construction doctrine in resolving workers’ compensation disputes. During our sample period from 1997 to 2016, four states passed new statutes prohibiting its use. A West Virginia amendment in

---

\(^{8}\) Ohio provision: R.C. § 4123.95; Maryland provision: MD LABOR EMPLY § 9-102

\(^{9}\) Town of Russell Volunteer Fire Dept. v. Labor and Industry Review Com’n, 223 Wis. 2d 723 (Ct. App. 1998)

\(^{10}\) West’s F.S.A. § 440.015: Legislative Intent

\(^{11}\) M.S.A. § 176.001: Intent of the Legislature
2003 stated:

“the Legislature hereby declares that any remedial component of the workers’ compensation laws is not to cause the workers’ compensation laws to receive liberal construction that alters in any way the proper weighing of evidence as required...”¹²

Similarly, Indiana legislation in 2006 implicitly prevented the adjudicator from making a presumption favorable to workers in issues related to burdens of proof.¹³

III.3 Intoxication Presumed As a Proximate Cause of Injury

In most of the states, intoxication is a statutory defense to reduce or eliminate compensation for an intoxicated worker injured on the job.¹⁴ A worker’s intoxication could significantly undermine the claim that the suffered injuries were caused by the underlying risk of the employment. Whether intoxication is presumed to be the proximate cause of work-related injuries has significant implications for the workers’ burdens of proof. If such presumption is not allowed, the employer bears the burden of establishing that intoxication was the proximate cause of the injury to avoid compensation liability. On the other hand, if such presumption is allowed, the injured worker bears the burden to show that their injuries are independent from the intoxication in order to receive compensation benefits. Allowing the presumption that intoxication is a proximate cause plausibly reduces the probability the injured worker will be compensated due to the added costs of collecting sufficient evidence to rebut the presumption.

A majority of the states disallow the presumption through common law precedents, statutes or a combination of both. The Minnesota workers’ compensation statutes explicitly require employers to show intoxication is the proximate cause of workers’ injuries.¹⁴ A Virginia decision held that an injured worker, who suffered fractures due to a fall from a forklift pallet while clearly intoxicated, was still entitled to workers’ compensation benefits.¹⁵ The ruling stated that the proximate cause of the injury was the backward movement of the forklift before lowering the pallet and the employer failed to establish that intoxication was the proximate cause of his injuries.

Prior to 1997, there were seven states that allowed the presumption of workers’ intoxication as

¹²WV ST § 23-1-1
¹³IC 22-3-7-2: “The burden of proof is on the employee. The proof by the employee of an element of a claim does not create a presumption in favor of the employee with regard to another element of the claim.”
¹⁴M.S.A. § 176.021: “If...the intoxication of the employee is the proximate cause of the injury, then the employer is not liable for compensation. The burden of proof of these facts is upon the employer.”
the proximate cause of injuries. During the sample period from 1997 to 2016, six additional states added new language allowing such a presumption.\textsuperscript{16} The new language in Missouri in 2006 stated:

“The voluntary use of alcohol to the percentage of blood alcohol sufficient under Missouri law to constitute legal intoxication shall give rise to a rebuttable presumption that the voluntary use of alcohol under such circumstances was the proximate cause of the injury. A preponderance of the evidence standard shall apply to rebut such presumption.”\textsuperscript{17}

III.4 Benefits Apportionment Due to Preexisting Conditions

Preexisting conditions can potentially prevent injured workers from recovering the full amount of compensation benefits. Workers with preexisting health conditions are typically eligible for compensation benefits only if the job-related injuries significantly aggravate the intensity of the preexisting conditions. Injured workers are not eligible for compensation benefits if their symptoms are a “mere natural progression” of their preexisting conditions.\textsuperscript{18}

Once “significant aggravation” of preexisting conditions is established, states have adopted one of two compensation regimes. About half of the states pay workers the full amount, which raises the financial burden on employers. To mitigate the financial burdens and to provide more employment opportunities to workers with preexisting conditions, some states have established “second injury funds” which partially cover compensation benefits payments to workers with preexisting conditions.

The rest of the states pay workers the fraction of benefits that matches the fraction of the condition that was solely induced by the workplace accident. These states usually only apportion permanent disability benefits. Workers can usually recover the full amount of temporary disability benefits, regardless of the status of their preexisting conditions.\textsuperscript{19}

To some extent the apportioning of benefits substantively relates directly to the amount of benefits paid. At the same time the apportionment also relates to burdens of proof because workers...
have to show that their medical symptoms are 100 percent attributable to the workplace accident in order to receive full benefits, and this often is difficult. Before 1997, there were 21 states with laws that apportioned benefits when workers had preexisting conditions. Over the sample period, three more states joined the benefits apportionment list. In California, the new language in 2004 read:

“(a) The employer shall only be liable for the percentage of permanent disability directly caused by the injury arising out of and occurring in the course of employment. (b) If the applicant has received a prior award of permanent disability, it shall be conclusively presumed that the prior permanent disability exists at the time of any subsequent industrial injury. This presumption is a presumption affecting the burden of proof.”

III.5 Medical-Care Restriction Laws

The rapid increases in medical costs over the last 40 years have directly affected the medical costs of treating workplace injuries under workers’ compensation. As a result, states have explored a variety of ways to control the costs while trying to maintain appropriate medical treatment of injuries. The two most common forms of medical legislation are laws that alter medical fee schedules and laws that establish state-wide networks of physicians to treat work-related injuries.

III.5.1 Medical Fee Schedule Laws

Adoption of a medical fee schedule puts a limit on the workers’ compensation payment to medical care providers for providing injured workers with various medical services. The fee schedule is typically adopted by the state agency in charge of administering the workers’ compensation program.

In states where the medical fee schedule system does not exist, there are other types of restrictions on medical benefit payments. Georgia state law requires the State Board of Workers’ Compensation to approve, on a case-by-case basis, the payment of physician fees, hospital charges, charges of prescription drugs or other costs of regulated medical items. All of the filed medical

---

20 California(2004), Kansas(2011) and Missouri(2006)
21 West’s Ann.Cal.Labor Code § 4664 (a) - (b)
22 It’s usually the commissioners of state workers’ compensation commission. But there are exceptions. For example, in Oregon, it is the Director of Consumer and Business Service who is in charge of promulgating rules for developing and publishing fee schedules for medical services.
payments are reviewed on the “reasonableness” standard by a panel of experts composed of peer physicians and hospital representatives.\textsuperscript{23}

Prior to 1997, there were 24 states that had adopted fee schedule systems for medical service. During our sample period from 1997 to 2016, an additional 15 states adopted fee schedule provisions for medical service payments. For example, Idaho adopted a fee schedule in 2005 by adding the following language:

“The commission may adopt a fee schedule to determine the allowable payments to be made to medical providers under this chapter, including but not limited to, the fee schedule the commission has adopted to determine the allowable payments to be made to medical providers under the Idaho worker’s compensation law.”\textsuperscript{24}

III.5.2 Physician Choice and Physician Network Laws

Choice of physicians is an important issue in the state workers’ compensation programs. Employers and workers have competing self-interests in determining providers of medical care. Employers want to treat workers’ injuries in a way that minimizes the medical costs. In contrast, if the medical services are covered, workers tend to have their injuries treated with the “best-quality” medical care and are less concerned about the costs. State laws developed various systems to account for the legitimate interests of both employers and workers.

In some states workers play a major role in determining medical care although their choices are usually subject to some restrictions.\textsuperscript{25} In other states, medical care is mainly determined by employers. State laws usually require employers to furnish medical care and workers’ choices are usually limited to what employers offer.\textsuperscript{26}

Physician networks normally consist of doctors specialized in treating occupational diseases or general fields of medicine.\textsuperscript{27} To join the network, medical care providers typically need to apply and get approved by the state agency in charge of administering the workers’ compensation programs. Employers may choose to join the network for treating workers’ work-related injuries, but this is

\textsuperscript{23}§ 34-9-205 (a)-(b)
\textsuperscript{24}ID ST § 72-1026
\textsuperscript{25}For example, in Hawaii, injured workers are free to choose any physician or surgeon who practice on the island. However, workers need to comply with fee schedules and provide sufficient notice to employers about the choice of physicians. See HI ST § 386-21.
\textsuperscript{26}In these states workers can only determine treating physicians under extraordinary circumstances, such as an emergency (See IN ST 22-3-5-4 and IA ST § 85.27) and an employers’ failure to provide medical care (ID ST § 72-432).
\textsuperscript{27}The physician network in California has the following website: \url{https://www.dir.ca.gov/dwc/mpn/dwc_mpn_faq.html}
not a mandate. If the employers are in the network, the physician choices of injured workers are constrained to choose a physician in the approved network.

Our analysis focuses on assessing the impacts of state laws that initiate and regulate physician networks to treat work-related injuries. Prior to 1997, three states had authorized and regulated physician networks. Over our sample period from 1997 to 2016, four states passed new laws that established such networks. For example, California passed the following provision in 2005:

“... an insurer, employer, or entity that provides physician network services may establish or modify a medical provider network for the provision of medical treatment to injured employees. The network shall include physicians primarily engaged in the treatment of occupational injuries. The administrative director shall encourage the integration of occupational and nonoccupational providers...”

IV Measuring the Monetary Values of Changes in Burden of Proof in Liability Settings

The total benefits paid to injured workers for wage replace can be described by the following equation:

\[ B_{it} = A_{it} \times s_{it} \times p_{it}(L_{it}) \times c_{it}(L'_{it}) \] (1)

\[ B_{it} \] is the total amount of benefits paid in state \( i \) and year \( t \); \( A_{it} \) is the number of workers injured or killed in workplace accidents; \( s_{it} \) is the statutory benefits to be paid for the injury; \( p_{it}(L_{it}) \) is the probability of receiving workers’ compensation benefits for an injury because it occurred in the course or out of employment; and \( c_{it}(L'_{it}) \) is the compensable share of the benefits to be paid for the injury, which is based on the added damage caused by the workplace accident above and beyond the worker’s preexisting injury status. \( L_{it} \) and \( L'_{it} \) are vectors of dummy variables of burden of proof laws that respectively influence \( p_{it} \) and \( c_{it} \). As examples, an increase in the burden of proof

28 Private insurers also establish physician networks without the support of state regulation. For example, “Omnet Gold” (OG) is a physician network for work-related injuries in Louisiana and was developed entirely by a for-profit private insurer, Louisiana Workers’ Compensation Corporation (LWCC), with little involvement from the state bureaucracy. Bernacki et al. (2006) performed a study that compared workers’ compensation claims managed by the OG health-care providers to claims managed by non-OG healthcare providers with the following results: (1) the average number of reported lost days for closed OG claims is 53, lower than non-OG claims’, 99; (2) the average medical care costs for closed OG claims is 5855, significantly lower than non-OG claims’, 9850 and (3) For closed OG claims, the average indemnity costs is 4864, lower than the non-OG claims’, 7881.

29 CA LABOR § 4616
that injury arose out of or in the course of employment would lower \( p_{it} \), as would rules related to compensation when the worker might have been impaired by alcohol consumption, while rules about preexisting injury status would influence \( c_{it} \).

Dividing both sides of Equation (1) by the total number of workers covered by the workers’ compensation programs, \( N_{it} \).

\[
\frac{B_{it}}{N_{it}} = \frac{A_{it}}{N_{it}} * s_{it} * p_{it}(L_{it}) * c_{it}(L'_{it})
\]

Replace \( \frac{B_{it}}{N_{it}} \) with \( b_{it} \), the benefits paid per covered worker, and \( \frac{A_{it}}{N_{it}} \) with \( a_{it} \), the annual accident rate for the worker.

\[
b_{it} = a_{it} * s_{it} * p_{it}(L_{it}) * c_{it}(L'_{it})
\]

We have information on \( b_{it} \) from the National Academy of Social Insurance. State-level data on nonfatal accidents is not reported for every state in every year. As a proxy for \( a_{it} \), we use the fatal accident rate \( f_{it} \). Since it is an imperfect proxy, we substitute

\[
\alpha_{it} = f_{it}^{\alpha} * \epsilon_{it}^{f}
\]

where \( \epsilon_{it}^{f} \) is an error term. To determine the statutory benefits \( s_{it} \), we created an expected benefit index that provides estimates of the present value of the statutory benefits for four different types of injury outcomes (fatal \((f)\), permanent total \((pt)\), permanent hand loss \((h)\), and temporary total \((tt)\)) based on the national average wage \( w_t \). We then develop an expected benefits index \( EB_{it} \) by multiplying the probabilities of each type of accident by the benefit for that type and then adding the four products together.

\[
EB_{it} = \alpha^f * s^f_{it}(w_t) + \alpha^{pt} * s^{pt}_{it}(w_t) + \alpha^h * s^h_{it}(w_t) + \alpha^{tt} * s^{tt}_{it}(w_t)
\]

Section 2 has more detailed description of how we construct the index. The expected benefit index is an imperfect measure of the actual statutory benefits for the accidents, so we multiplied the expected benefits index with an error term, \( \epsilon_{it}^s \).

\[
s_{it} = EB_{it}^{|s} * \epsilon_{it}^{s}
\]
We do not know the actual probabilities \( p_{it}(L_{it}) \) and \( c_{it}(L'_{it}) \). But we do know that the burden of proof laws affect these probabilities. We make the following assumptions:

\[
p_{it}(L_{it}) = \varepsilon_{it}^p \ast \exp(L_{it})^{-\lambda} \tag{7}
\]

and

\[
c_{it}(L'_{it}) = \varepsilon_{it}^c \ast \exp(L'_{it})^{-\theta} \tag{8}
\]

where \( \varepsilon_{it}^p \) and \( \varepsilon_{it}^c \) are error terms that capture other unmeasured aspects of the probabilities.

After substituting Equations 4, 6, 7, and 8 into Equation 3, the equation becomes as follows:

\[
b_{it} = f_{it}^\alpha \ast EB_{it}^\beta \ast \exp(L_{it})^{-\lambda} \ast \exp(L'_{it})^{-\theta} \ast \varepsilon_{it}^f \ast \varepsilon_{it}^s \ast \varepsilon_{it}^p \ast \varepsilon_{it}^c \tag{9}
\]

After taking the natural log on both sides of the equation, it becomes

\[
\ln b_{it} = \alpha \ln f_{it} + \beta \ln EB_{it} + \lambda L_{it} + \theta L'_{it} + \ln \varepsilon_{it}^f + \ln \varepsilon_{it}^s + \ln \varepsilon_{it}^p + \ln \varepsilon_{it}^c \tag{10}
\]

Let \( \ln \varepsilon_{it} = \ln \varepsilon_{it}^f + \ln \varepsilon_{it}^s + \ln \varepsilon_{it}^p + \ln \varepsilon_{it}^c \), the final estimation equation becomes

\[
\ln b_{it} = \alpha \ln f_{it} + \beta \ln EB_{it} + (\lambda + \theta) L_{it} + \ln \varepsilon_{it} \tag{11}
\]

where the coefficients \( \alpha \) and \( \beta \) are the elasticities of the benefit payments per covered worker with respect to the fatal accident rate \( f_{it} \) and the expected benefits \( EB_{it} \), respectively. \( \lambda \) and \( \theta \) measure the relationship between the benefits payments per covered worker and the presence of different types of burden of proof laws. Section VI proposes an econometric framework of using generalized difference-in-difference econometric models to identify the causal effects of the burden of proof laws on benefits payments.

V Data

We construct a state-by-year panel dataset from 1997 to 2016 by putting together data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states.
V.1 Cash Benefits and Medical Benefits

The dependent variables in the analysis are the cash benefits per covered worker and medical benefits per covered worker. The cash benefits are the payments to replace lost earnings for injured or deceased workplace accident victims. Covered workers refer to workers on employer payrolls who are eligible to receive benefits when injured. The information comes from annual NASI reports titled *Workers’ Compensation Benefits, Coverage, and Costs* from 1997 through 2016. To obtain measures in 2010 U.S. dollars, we deflate the cash benefits by the national Consumer Price Index for urban areas and medical benefits by the national Consumer Price Index for medical care. Figure 1.1 shows the trends of these calculations at the national level.

V.2 Workers’ Compensation Expected Benefits Index

To control for statutory changes in the payments for different types of accidents we follow Allen (2015) and Kantor and Fishback (1995) in creating an “expected benefits index.” The state laws generally stipulate different payment schemes for various accident outcomes. We divided injuries into four types: death, permanent total disability, permanent partial disability, and temporary total disability. With a few exceptions, the benefits payment schedules specify a weekly payment duration and a weekly payment amount that is a percentage of usual weekly earnings up to a weekly maximum and not below a weekly minimum. Most state laws index the weekly maximums to the annual state-level average weekly wage. Weekly cash benefits are determined by multiplying workers’ weekly wage by a replacement percentage that is specified in statutes, usually 2/3. The maximum weekly wage replacement data are collected from the Social Security Administration as well as state statutes.

For each accident type we calculate the present value of the stream of weekly payments using a discount rate of 5 percent and the national average weekly wage for that year. We chose the national average weekly wage to ensure that the variation in the payments across states was driven only by statutory differences in the state laws. We then calculated real present values in 2010 dollars.

---

30 The reports can be downloaded from the following link: https://www.nasi.org/research/workers-compensation

31 The Consumer Price Index data comes from the US Bureau of Labor Statistics and can be downloaded from the following website: https://www.bls.gov/cpi/data.htm

32 The SSA maintains and regularly updates a chart of “States’ Maximum Workers’ Compensation(WC) Benefits, which is publicly available and could be found from the following link: https://secure.ssa.gov/poms.nsf/lnx/0452150045

33 SSA develops a national average wage index for benefits computation purpose. Our national average wage is calculated by dividing the index by 52. The index data can be found in the following website: https://www.ssa.gov/oact/cola/AWI.html
using the Consumer Price Index. Then, we calculated the weighted sum of the real present values of the four types of accident payments with the accident probabilities as weights.\(^{34}\) The expected benefit index offers a single value for the generosity of workers’ compensation cash benefits that incorporates information on the generosity for all types of accidents. We held the accident rate weights constant across time so that the only reasons for the changes in the state expected benefits is differences in statutes and changes over time in the national average wage. Figure 1.2 shows that the national average of these real expected benefit payments has a strong upward trend. To get the downward trend in national average cash payments per covered workers in Figure 1.1 there must be factors causing a strong downward trend that more than offset the rising trend in benefits.

We developed this expected benefits index primarily to address the Pro Publica survey’s imprecise measures of changes in state-level compensation benefits. Pro Publica’s measure is problematic because it does not change from year to year the large majority of the US states that index their compensation benefits with the state-level average weekly wages or inflation. In those states, the Pro Publica measure registers that the states index their benefits, but it would only change if the state changed its indexing formula.

To illustrate the problem, consider the change in the level of nominal benefits in states without and without indexing. In Georgia, which did not index benefits, Pro Publica’s measure correctly records increases in maximum weekly benefits when Georgia raised those benefits in the years 2005, 2007, and 2013. Their measure incorrectly shows no change in benefits in the states that indexed, even though the maximum weekly payments automatically increased every year without new legal amendments. Thus, Pro Publica’s measure misses the rise in nominal maximum benefits every year for a large majority of states.\(^{35}\) In fact, after adjusting for inflation, we found that the real weekly maximum benefits in many states rose over the period. Pro Publica’s measure is also problematic for comparing changes across states over time in the ratio of maximum benefits to weekly wages. Given the rise in wages and inflation, Georgia’s benefit ratio actually fell in nearly every year except 2005, 2007, and 2013, while the Pro Publica’s measure for Georgia incorrectly stayed constant.

Our expected benefits index incorporates the specific amounts of weekly maximum benefits in

\(^{34}\)The accident probability data come from Accident Facts, which was published by the National Safety Council in 1997. At the national level, the share of workers’ compensation claims that involved permanent partial disabilities rose from 27.7 percent in 1997 to 36.9 percent in 2016, which would have contributed to an increase in benefits per worker. If the states generally followed the national trend, the year fixed effects would have been effective controls. If the states had different trends, the state-specific time trends would have been effective controls.

\(^{35}\)According to Allen (2004), as of 2000, all states except Alaska, Arizona, California, Georgia, Indiana, Minnesota, and New York were indexing their maximum weekly benefits for temporary total disability to an average wage or inflation rate. Indexing was then adopted by Arizona and New York in 2007 and reintroduced by California in 2007 and Minnesota in 2013.
both “indexing” and “non-indexing” states, which more accurately reflect the state-level changes in statutory benefits than the *Pro Publica* measures. In addition, our index also reflects legislative changes in other parameters that affect the payments of workers’ compensation benefits, such as changes in death benefits and the payment duration for certain types of disabilities.\(^{36}\) Therefore, this index aggregates changes in statutory benefits in a way that shows what workers might expect and it also allows us to better determine which types of laws contributed to the decline in benefits per covered worker and which did not.

### V.3 Fatal Accident Rate

One of the factors that has contributed to the decline in benefit payments per covered worker in Figure 1.1 is the improvement of workplace safety, reflected by a decline in accident rate. Information by state and year for the number of fatal accidents and the number employed in private industry come from the Bureau of Labor Statistics (BLS).\(^{37}\) We then calculate the number of fatal workplace accidents per 100,000 workers in private industry for each year and state. Fatal injury information is available in all of the states as reporting of workplace fatal injuries is mandatory. But non-fatal injury information is only available in states that are willing to participate in the BLS Survey of Occupational Injuries and Illness, which is not federally mandated. Due to substantial number of missing values for the non-fatal injury data, we chose to use the fatal accident data as the proxy of workplace safety environment. The national average of the fatal accident rate in Figure 1.3 fell by about one-third between 1997 and 2016. This decline is potentially a reason for the declines in benefits per covered worker seen in Figure 1.1. The national average for missed days due to nonfatal injuries in the reporting states also showed a significant decline over the study period.

---

\(^{36}\) *Pro Publica* reported 27 legislative changes in death benefits, all of which are incorporated in this index measure. Our benefit measure also incorporates more than a dozen of recorded legislative changes in benefit payments duration. For example, West Virginia passed a law in 2003 which imposed a 2-year maximum duration for benefit payments to temporarily disabled workers regardless of whether they recover from the injuries.

\(^{37}\) BLS maintains and regularly updates data on state-level occupational injuries, illness and fatalities. The data can be access using the following link: [https://www.bls.gov/iif/state_archive.htm#AZ](https://www.bls.gov/iif/state_archive.htm#AZ) The employment data come from the Current Employment Statistics (CES) in BLS. The data are the seasonally adjusted number of employment in private industry at given state and year. The data can be assessed from the following link: [https://www.bls.gov/data/#employment](https://www.bls.gov/data/#employment)
V.4 State Laws

The set of state laws collected by Pro Publica inspired our interest in studying the issue, but Qin (2015) only collected new state laws between 2002 and 2014.\footnote{Their full study is at the following link: https://projects.propublica.org/graphics/workers-comp-reform-by-state} We expanded on their work by using the statute volumes for the states from WestLaw to establish the status quo as of 1997 and to add laws passed in other years through 2016. For each of the category of laws not directly related to benefit levels, we created dummy variables with a value of one if the law was present in the state and zero otherwise. We also developed a law index which keeps track of state laws that increased workers’ burdens of proof to receive compensation benefits. As shown in the discussion of the state laws, some had versions of a law throughout the time period, others added the law during the time frame and others did not. Table 1.1 summarizes the states and years when each types of laws were passed. In the empirical analysis, the combinations of these state laws and the expected benefits index covers a substantial majority of the laws incorporated in the Qin (2015).

VI Econometric Strategy: Generalized Difference-in-Difference Models

Our goal is to examine how the laws influence the total benefits paid out per covered worker, which is essentially every worker in the sectors covered by the workers’ compensation law. Thus, covered workers include the non-injured, the injured who receive benefits, and the injured who do not receive benefits. Scholars and practitioners predict that each of the five laws would reduce average benefits per covered worker because fewer injured workers would obtain benefits. However, it is possible that a higher burden of proof might also result in “selection effects” leading to an increase in the average benefits paid per claim to injured workers who actually receive benefits. This would happen if some injured workers decide not file claims because the costs of meeting the burden of proof are greater than the expected benefits payments that injured workers receive.\footnote{Priest and Klein (1984) developed an economic model of litigation which imply that changes in legal standards affect quality of disputes that individuals decide to litigate. Empirically, scholars studied the case quality selection effects of new legal standards by evaluating the causal impact of the heightened pleading standard introduced by Bell Atlantic v. Twombly, 550 U.S. 544 (2007) and Ashcroft v. Iqbal, 129 S.Ct. 1937 (2009). Twombly and Iqbal replaced the previous “no-set-of-fact” pleading standard with a new “plausibility” standard, which raised the evidentiary bar for litigated cases to survive a motion to dismiss. In empirical studies the new standard had little impact on the probability of granting motions to dismiss (Hubbard 2013, 2017) or of surviving summary judgment (Gelbach 2016). Gelbach (2011) developed an economic model that outlines the channels through which higher quality cases are selected into litigation through decisions by defendants and plaintiffs in the settlement process.}
Our empirical analysis examines the final effects of the laws on cash and medical benefit payments per covered worker.

To identify the causal effects of the laws on average benefit payments per covered worker, we start with Equation (11) and employ the following generalized difference-in-difference model, or two-way fixed effects model, as the main specification:

$$\text{Log Benefits}_{st} = \beta \text{Legislation}_{st} + \delta_s + \tau_t + X_{st}\gamma + \theta_s t + \varepsilon_{st}$$  \hspace{1cm} (12)

The outcome variable is \(\text{Log Benefit}_{st}\), the natural logarithm of benefit payments per covered worker, either cash or medical, in state \(s\) at year \(t\). \(\delta_s\) is a vector of state fixed effects, controlling for time-invariant and state-specific unobserved characteristics, such as workers’ demographics and industry composition. \(\tau_t\) is a vector of year fixed effects, accounting for national shocks like changes in federal OSHA regulations. \(X_{st}\) includes other factors discussed in Equation (11): (1) \(\text{Log Accident Rate}_{st}\), the natural logarithm of the fatal accident rate in state \(s\) and year \(t\); (2) \(\text{Log EB Index}_{st}\), the natural logarithm of workers’ compensation expected benefits index in state \(s\) at year \(t\). The expected benefits index variable is only included for analyzing cash benefits, as it does not capture any information related to medical benefit payments. We also control for a vector of state-specific time trends, \(\theta_s t\), as a robustness check. Standard errors are clustered at the state level (Bertrand et al., 2004).

\(\text{Legislation}_{st}\) represents various measures of changes in workers’ compensation laws. First, we create dummy variables for the five types of state laws. Each dummy has a value of one if the law is in place and zero otherwise. Thus states with the relevant law as of 1997 have a value of one throughout the sample period. In addition, we create a law index which aggregates three types of legislative changes that increase workers’ burdens of proof. The law index value increases by one when state \(s\) in year \(t\) passed one of the following three types of laws: laws that ban the liberal construction of statutes, laws that presume intoxication as the proximate cause of injury, and laws that apportion the workers’ compensation benefits due to preexisting conditions. The index also accounts for the number of burden of proof laws that existed at state level prior to 1997.

\(\beta\) is the key parameter vector of interest, which captures the percentage change in benefits payments per worker resulting from the burden of proof and the medical workers’ compensation laws and is the sum of \(\lambda\) and \(\theta\) parameters in Equation (11) that capture the impact of the laws on the probability of benefit payments and the compensable share of benefits, respectively. Our
identification exploits the cross-state and cross-time variation in the passage of the new workers’ compensation laws. To interpret $\beta$ as causal estimates, we make the following two key assumptions: (1) conditional on observable controls, state and year fixed effects, the law dummies are uncorrelated with the error term, $\varepsilon_{st}$; (2) during the pre-treatment periods, the benefits payments in treatment and control states follow parallel trends. Assumption (1) might be violated if there is endogeneity bias based on reverse causation related to the political economy of the laws. But we anticipate in each case that it will be positive. States are more likely to adopt measures to limit benefits of both kinds in settings where benefits are high and rising faster than anticipated. If this prediction holds, the coefficient estimates would understate the impact of the laws on benefit payments per covered worker if the estimates of $\beta$ have negative signs.

To test the parallel pre-trend assumption, we estimate the following event study model:

$$Log\ Benefits_{st} = \sum_{i=-T}^{T} \lambda_i Legislation_{st+i} + \delta_s + \tau_t + X_{st}\gamma + \theta_s t + \varepsilon_{st}$$  \hspace{1cm} (13)$$

Compared with Equation (12), the event study model decomposes law variables into a vector of leads and lags of the laws. We create four lags and leads for each measure of workers’ compensation laws. $i$ indexes the lags or leads and take integer values from -3 to 4. The main coefficients of interests are a vector of $\lambda_i$. The estimates of leading terms of $\lambda_i$ provide diagnostic tests on the validity of the parallel trend assumption. $\lambda_0$ measures the instantaneous causal effects of the laws on benefit payments. The lagged terms of $\lambda$ capture the dynamic treatment effects of the laws on the outcomes.

The recent econometrics literature suggests that two-way fixed effect models produce biased estimates of average treatment effects if the treatment effects are not homogeneous over time (Goodman-Bacon 2018; Athey and Imbens 2018; De Chaisemartin and d’Haultfoeuille 2020). Goodman-Bacon (2018) demonstrates that the two-way fixed effects estimator is a weighted average of all possible 2 by 2 difference-in-difference estimators that compare the timing of treatment. Using earlier-treated groups as “controls” to estimate the treatment effects of later-treated groups causes an issue of “negative weights” if the treatment effects are time-varying (Goodman-Bacon 2018; De Chaisemartin and d’Haultfoeuille 2020). Time-varying treatment effects also contaminate the event study estimators (Borusyak and Jaravel 2017; Sun and Abraham 2020; de Chaisemartin and D’Haultfoeuille 2020). Sun and Abraham (2020) shows that event study estimators of contemporaneous and dynamic treatment effects, first used in Autor (2003), are invalid if the treatment
effects are heterogeneous across different groups and time.

To address the concerns over time-varying treatment effects, we estimate the weights associated with $\beta$ in Equation (12), which provide helpful diagnostics of whether two-way fixed effects estimator has the opposite sign to the average treatment effects. In addition, we estimate the event study model, Equation (13), using the $DID_M$ estimators proposed in De Chaisemartin and d’Haultfoeuille (2020 A & B) to show the evolutionary path of the treatment effects. Under this estimation framework, the estimates of the leading $\lambda$ terms are the placebo treatment effects that test the validity of parallel trend assumption. We anticipate placebo treatment effects to be small and statistically insignificant if the parallel trend assumption holds. The lagged treatment effects are estimated by using all the groups that are not treated at that time as “controls,” which ensures that the heterogeneous treatment effects across time do not bias the estimates of the lagged effects.

VII Results for the Effects of Burden of Proof Law Index

Table 1.2 shows the OLS estimates of how cash benefit payments are affected by the burden of proof law index when various types of fixed effects are added. The preferred specification is in Column (3) which include control variables, state fixed effects, and year fixed effects. The Column (4) specification restricts the sample to a subset of states that had not passed any type of legislation captured in the burden of proof law index prior to 1997. Specification (5) further includes state-specific time trends as a robustness check. The point estimate in Column (3) for the burden of proof law index is -0.2 and is statistically significant at the 1% level. This result suggests that cash benefit payments per covered worker were 20 percent low after states passed laws that increase workers’ burdens of proof. The coefficient changes to -0.26 when the sample is restricted to states that had no burden of proof laws prior to 1997. The coefficient in Column (5) is similar and suggests that the result is also robust to including state-specific time trend into the model.

To better understand the treatment effects, we computed the weights associated with the estimate of the Burden of Proof Law Index. 52% of the weights are negative, which, according to Goodman-Bacon (2018), implies that the treatment effects are dynamic. The negative weights

---

40 The weights are estimated using the stata package, TWOWAYWEIGHTS. De Chaisemartin and d’Haultfoeuille (2020) suggests that this stata package produces valid estimates of weights for both binary and non-binary treatment variables under the two-way fixed effects model framework.

41 According to De Chaisemartin and d’Haultfoeuille (2020 A & B), the $DID_M$ estimators apply to both binary and non-binary discrete treatment variables. Therefore, the method can easily accommodate the fact that the burden of proof law index is a non-binary, discrete variable. We implement the estimation methods using the stata package, DID_MULTIPLEGT, which was developed by de Chaisemartin et al. (2019).
sum up to -0.91. Figure 1.4 shows the estimates of the event study estimation using the methods developed in de Chaisemartin and D’Haultfoeuille (2020). The treatment effects trajectory implies that the passage of the burden of proof laws leads to negative impacts that are quite different from the preexisting trends in cash benefit payments. The estimated coefficients of the pre-treatment periods have small magnitude. The estimate of one-year lead has a positive sign and is nearly statistically significant at the 5 percent level. Such slight upward bias tends to underestimate the negative treatment effects of the law. Therefore, the estimated coefficient for the law index might be interpreted as a lower bound of the treatment effects. The estimated coefficients for the post-periods imply that the effects of the laws in reducing cash benefit payments become stronger over time. The laws decrease the cash benefits by about 11% in the first year after the passage. The negative effects increased to about 24% in the third year after the laws being passed.

Another key finding in Table 1.2 is that there is a strong positive relationship between the cash benefit payments and expected benefits index. As both the payment and the index are in natural logarithm forms, the estimated coefficient is the elasticity of actual benefit payments per covered worker with respect to statutory benefits, which is similar to the “benefit elasticity” in the literature. In Column (3) of Table 1.2, the point estimate of the expected benefits index (log) is 1.8 and statistically significant at the 1% level. The estimate further decreases to 1.18 in Column (5) after including state-specific time trends into the model. This result suggests that a 1 percent increase in the expected benefits index is associated with a 1.18 percent increase in the cash benefit payments. The estimated elasticity is slightly larger than 1 and consistent with the estimate by Krueger and Burton Jr (1990) where they found a one unit increase in workers’ compensation benefits is approximately associated with a one unit increase of employers’ costs in paying workers’ compensation claims. Finally, although we expected a positive coefficient for the fatal accident rate, the elasticity is very small at -0.007 and statistically insignificant.

Table 1.3 shows the OLS estimates of how medical benefit payments are affected by the burden of proof law index. Column (3) is the preferred specification. The estimated coefficient for the burden of proof law index is -0.157 and it is statistically significant at the 1% level. The result implies that the passage of laws that increase workers’ burdens of proof decreased the medical benefit payments by 15.7 percent, which has a slightly smaller magnitude than the effects of the law index on cash benefit payments. About 53% of the estimated weights are negative and they sum up to -0.91. Also, the coefficients in Column (4) and (5) suggest the result is robust to using a sub-sample of states that did not have the laws prior to 1997 and including state-specific time
trends in the regression. Figure 1.5 shows the evolutionary path of the treatment effects of laws on medical benefit payments. The coefficients of the pre-treatment periods are close to zero and not statistically significant and suggest that the medical benefit payments in treatment and control states were trending similarly prior to the passage of the laws, which supports the parallel trends identifying assumption. The estimates of the post-treatment periods suggest that the negative effects of the laws in reducing medical benefit payouts did not kick in until the second year after the passage of the laws. The magnitude of the effects is around negative 17%.

The next step is to address how much of the decline in the national average of benefit payments per covered worker between 1997 and 2016 was associated with the new burden of proof laws. Using the means for the variables for 1997 and 2016 and the coefficients, we developed a decomposition analysis in Table 1.4 to see how much of the drop in benefits per covered worker are associated with the burden of proof law index (Oaxaca and Ransom 1994). During the sample period, the means of natural log of cash and medical benefit payments per covered worker fell by -0.37 and -0.29, respectively. The decomposition in Column (5) shows the change in mean value in the burden of proof law index times its coefficients in Column (3) specification in Table 1.2 and 1.3 as a percentage of the changes in log cash and medical benefits per covered worker. For example, the share of states with any type of the burden of proof laws aggregated in the index rose by 0.2 from 0.7 in 1997 to 0.9 in 1999. The 0.2 rise is then multiplied the coefficient -0.201 from Column (3) of Table 1.2 to find that the rise in burden of proof laws are associated with a reduction of -0.0402 in the log of cash benefit payments per covered worker. The burden of proof laws therefore account for 10.9 percent of the -0.37 drop in log cash benefit payments per covered worker. Using a similar analysis, we find that the burden of proof law index is associated with 10.8 percent of the drop in medical benefit payment per covered worker.

VIII Results for Each Type of the Laws and Decomposition Analysis

We further investigate the effects of each type of workers’ compensation laws on cash and medical benefit payments. We estimate the causal effects of each type of the laws on cash and medical benefit payments using two-way fixed effects regressions and event study models. The analysis further clarifies the specific types of laws that drive the results from burden of proof law index. It is worth noting that different types of laws might work jointly in reducing workers’ compensation
benefit payments. Therefore, compared with type-specific law dummies, the burden of proof law index tends to capture a more wholistic picture of the impacts of the new laws. Appendix A provides more detailed results on the effects of each type of the laws. Appendix B presents the event study graphs, which test the validity of the parallel trends assumption and also show the evolutionary trajectory of the treatment effects.

Table 1.5 shows the OLS estimates of how cash benefit payments are affected by various workers’ compensation laws. The results suggest that cash benefit payments are lower in states that adopted these types of laws. Based on the Column (6) specification with all law dummies included, the cash benefits are 24.2 percent lower in states that prohibit the use of the literal construction rule, 13.5 percent lower in states with intoxication laws, 11.9 percent lower in states where benefits can be apportioned based on workers’ preexisting conditions, 5.7 percent higher in states that adopted medical fee schedules and 8.3 percent lower in states with physician network laws. The coefficient for the liberal construction ban is statistically significant at 5 percent level. Other law coefficients, compared with those from separate regressions, have smaller magnitudes and are not precisely estimated, suggesting that there was synergy of different types of laws in the law index that contributed to its negative impacts on the cash benefit payments.

Using the means for the variables in 1997 and 2016 and the coefficients from Column (6) in Table 1.5, Table 1.6 shows the results of decomposition analysis of how much the drop in cash benefit payments per covered worker is explained by the passage of the new workers’ compensations laws and changes in other variables. Column (5) in Table 1.6 shows that the liberal construction ban, the intoxication law, the apportionment laws, and the physician network laws contributed 5.2, 2.2, 1.3, and 1.4 percent of the drop in cash benefit payments. In total, the four types of workers’ compensation laws can explain 11.1 percent of the declines in the national average of cash benefit payments per covered worker. These results show that the laws had strong effects in the states where they were enacted, but their ability to explain the drop in the national average of log cash benefits per worker is relatively weak because only 8 to 24 percent of the states adopted the laws. Meanwhile, the decomposition for the expected benefits index, fatal injury rate, and medical fee schedule laws had negative signs because their trends might have served to raise cash benefits per covered worker and thus pushed against the -0.37 drop in the mean of the log cash benefits per covered worker.

For each type of the workers’ compensation laws, we use natural logarithm of cash and medical benefit payments as the outcome variables and estimate the same set of regression specifications as Table 1.2 and Table 1.3. Overall, the results are robust to different regression specifications.
Table 1.7 shows the OLS estimates of how medical benefits are affected by the workers’ compensation laws. The point estimates suggest that the medical benefit payments per covered worker are 9.5% lower in states that banned liberal constructions, 21% lower in states with new intoxication legislation, and 24% lower in states that adopted physician network laws. All of the point estimates are statistically significant. Apportionment laws and fee schedule laws do not have statistically significant relationships with medical benefit payments per covered worker. The result for medical fee schedule laws is consistent with Pozzebon (1994), which did not find evidence that having a medical fee schedule was negatively correlated with medical benefit payments. It is plausible that states, without formal medical fee schedules, might have other types of restrictions on reimbursement of medical care expenses, such as some form of administrative review.

Using the means for the variables in 1997 and 2016 and the coefficients from Column (6) in Table 1.7, the results of a decomposition analysis of the mean of the natural log in medical benefit payments per covered worker are shown in Table 1.8. The mean of the natural log of the medical benefits per covered worker fell by -0.29 from 5.42 in 1997 to 5.13 in 2016. The decomposition exercise suggests that the laws banning liberal construction can explain 2.6 percent of the drop in the national medical benefit payments, the intoxication laws can explain 4.3 percent of the drop, and physician network laws can explain 4.9 percent of the drop. In total, the three types of new workers’ compensation laws explain about 11.8 percent of the declines in the national average of medical benefit payments per covered worker. The decline of fatal accident rate explains about 0.4 percent of the drop in medical benefit payments. In addition, the decomposition of apportionment laws and medical fee schedule laws have negative signs and might have served to increase the medical benefit payments as the coefficients of these variables are not statistically significant.

Finally, we estimate event study models to test the validity of the parallel trend assumption and also to generate the evolutionary paths of the treatment effects. The results are in Appendix B. Overall, the event study results suggest that the negative effects of liberal construction bans and apportionment laws are not the results of preexisting trends and therefore more likely to be causal. In Figure 1.6.B, the coefficients of the pre-treatment periods for medical benefits suggest no pre-treatment trends. In Figure 1.6.A, prior to the liberal construction ban cash benefit payments in the treatment states trended slightly higher, the opposite direction of the increasingly negative effects of the law after its passage. Figure 1.7.A shows the validity of parallel trends assumption for cash benefits, followed by statistically significant and increasingly negative effects of the apportionment laws on cash benefits after they were passed. In Figure 1.7.B, there is no evidence suggesting
violations of parallel trends assumption. The period-specific coefficients are not precisely estimated, although the point estimates of the post-treatment periods were much lower than before the laws were passed.

In Figure 1.8A, the parallel trend assumption for the negative effects of physician network laws on cash benefit payments holds, while the coefficients for the post-treatment periods are negative but not precisely estimated. The results contrast with the downward relative trend in medical benefit payments for states that adopted the physician network laws in Figure 1.8B. The treatment effects after the laws were passed held steady around a statistically significant value of -0.22 from period 1 through 4. This result casts some doubts on the causal interpretation of the estimated effects of physician network laws on medical benefits. Figures 1.9 and 1.10 show the event study graphs for the effects of intoxication laws and medical fee schedule laws on cash and medical benefit payments. Although there does not appear to be evidence suggesting violations of the parallel trends assumption, the coefficients have large standard errors and the trajectories of the treatment effects have noisy patterns, particularly for the effects of medical fee schedule laws. Therefore, we are reluctant to claim that these two types of laws caused reduction in workers’ compensation cash or medical benefit payments per covered worker.

IX Conclusion

Analyzing the recent changes in workers’ compensation has offered a good opportunity to obtain clean estimates of the effect on benefit payments of burden of proof laws because fault is not an issue and statutes set the compensation level. The results here document that these types of laws have powerful effects on workers’ compensation benefits per covered worker. After some form of a workers’ compensation burden of proof law was introduced in a state between 1997 and 2016, real cash benefits and real medical benefits per covered worker fell by 20 and 16 percent, respectively.

Workers’ compensation scholars and members of the workers’ compensation community have suggested that the general trends in workers’ compensation laws have been reducing the benefits to paid to workers. Two sets of facts have fueled that view. First, the national average of benefits per covered worker fell between 1997 and 2016 shown in Figure 1.1. Second, Qiu (2015) developed measures of the changes in workers’ compensation laws for Pro Publica that suggested that the laws were becoming less favorable to workers. As part of this study, we separately estimated the impacts of the leading types of laws that they examined. The estimated results for the burden of
proof laws show that they served to reduce significantly the payments to workers. Decomposition suggests that the pattern of adoption of the burden of proof laws accounted for about 11 percent of the national decline in benefits per covered worker between 1997 and 2016. The physician network laws also served to reduce national medical payments per covered worker by about 4 percent.

Not all of the workers’ compensation changes during this period contributed to the decline. We calculate a measure of real expected benefit payments using the workers’ compensation laws that determine the benefit percentage of earnings, weekly maximum benefits, the length of time benefits will be paid, and other features that directly determine the wage replacement payments. The national average for that measure in Figure 1.2 has risen between 1997 and 2016 in large part because most states have indexed their maximum weekly benefits to inflation or to average earnings at the state or national level. The regression results show that the elasticity of benefits per covered worker with respect to the statutory expected benefit measure is around 1.2; therefore, the rise in real statutory benefits has been working against the decline in national benefits per covered workers.
Figures and Tables

Figure 1.1: National Cash and Medical Benefit Payments Per Covered Workers in 2010 Dollars

Notes: State-by-year level data of cash benefit payments, medical benefit payments and the number of covered workers come from the annual workers’ compensation reports issued by National Academy of Social Insurance (NASI). The nominal cash benefits are deflated by the urban Consumer Price Index (CPI) and the medical benefits are deflated by the medical CPI collected from Bureau of Labor Statistics (BLS).
Figure 1.2: National Workers’ Compensation Expected Cash Benefits Index in 2010 Dollars

Notes: State-level weekly maximum benefits and national level average wage index data come from Social Security Administration (SSA). The nominal benefits are deflated by CPI collected from BLS.
Figure 1.3: National Fatal Accident Rate

Notes: Fatal accident rate is defined as the number of fatal accidents per 100,000 workers in private industry. The measure is calculated with the following two strands of state-level data from BLS: seasonally adjusted employment and total counts of reported occupational fatalities.
Notes: The graph shows the point estimates and the 95% confidence intervals of the $\lambda$ terms in Equation (13). The outcome variable is the natural logarithm of cash benefit payments per covered worker. The treatment variable is an index of the laws that increase workers’ burdens of proof to receive compensation benefits. We estimate the coefficients by following de Chaisemartin and D’Haultfoeuille (2020). The standard errors are clustered at the state level. The sample period is from 1997 to 2016.
Figure 1.5: The Dynamic Effects of Burden of Proof Laws on Medical Benefit Payments

Notes: The graph shows the point estimates and the 95% confidence intervals of the $\lambda$ terms in Equation (13). The outcome variable is the natural logarithm of medical benefit payments per covered worker. The treatment variable is an index of the laws that increase workers’ burdens of proof to receive compensation benefits. We estimate the coefficients by following de Chaisemartin and D’Haultfoeuille (2020). The standard errors are clustered at the state level. The sample period is from 1997 to 2016.
Figure 1.6: The Effects of Liberal Construction Bans on Workers' Compensation Benefit Payments: Event Study Graphs

Notes: The outcome variables in Panel A and B are cash and medical benefit payments per covered worker, respectively. The graphs show the point estimates and the 95% confidence intervals of $\lambda_i$ terms in Equation (13). We estimate the coefficients by following de Chaisemartin and D'Haultfoeuille (2020). The standard errors are clustered at the state level. The sample period is from 1997 to 2016.
Figure 1.7: The Effects of Apportionment Laws on Workers’ Compensation Benefit Payments: Event Study Graphs

Notes: The outcome variables in Panel A and B are cash and medical benefit payments per covered worker, respectively. The graphs show the point estimates and the 95% confidence intervals of $\lambda_i$ terms in Equation (13). We estimate the coefficients by following de Chaisemartin and D’Haultfouille (2020). The standard errors are clustered at the state level. The sample period is from 1997 to 2016.
Figure 1.8: The Effects of Physician Network on Workers’ Compensation Benefit Payments: Event Study Graphs

Notes: The outcome variables in Panel A and B are cash and medical benefit payments per covered worker, respectively. The graphs show the point estimates and the 95% confidence intervals of $\lambda_i$ terms in Equation (13). We estimate the coefficients by following de Chaisemartin and D’Haultfœuille (2020). The standard errors are clustered at the state level. The sample period is from 1997 to 2016.
Figure 1.9: The Effects of Intoxication Laws on Workers’ Compensation Benefit Payments: Event Study Graphs

Notes: The outcome variables in Panel A and B are cash and medical benefit payments per covered worker, respectively. The graphs show the point estimates and the 95% confidence intervals of $\lambda_i$ terms in Equation (13). We estimate the coefficients by following de Chaisemartin and D’Haultfouillée (2020). The standard errors are clustered at the state level. The sample period is from 1997 to 2016.
Figure 1.10: The Effects of Fee Schedule Laws on Workers’ Compensation Benefit Payments: Event Study Graphs

Notes: The outcome variables in Panel A and B are cash and medical benefit payments per covered worker, respectively. The graphs show the point estimates and the 95% confidence intervals of $\lambda_i$ terms in Equation (13). We estimate the coefficients by following de Chaisemartin and D’Haultfoeuillé (2020). The standard errors are clustered at the state level. The sample period is from 1997 to 2016.
Table 1.1: Timing of States’ Passages of Workers Compensation Laws

<table>
<thead>
<tr>
<th>Category</th>
<th>Existed Prior to 1997</th>
<th>Passed Between 1997 and 2016</th>
</tr>
</thead>
</table>

Source: This summary is based on ProPublica publication of state-level workers’ compensation reform and our own research using the Westlaw database. The ProPublica publication reform can be assessed in the following link: [https://projects.propublica.org/graphics/workers-comp-reform-by-state](https://projects.propublica.org/graphics/workers-comp-reform-by-state). Westlaw consists of the workers’ compensation statutes for all of the states.
Table 1.2: The Effects of Burden of Proof Laws on Cash Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Benefits Index (log)</td>
<td>0.139</td>
<td>-1.329</td>
<td>1.792</td>
<td>1.340</td>
<td>1.178</td>
</tr>
<tr>
<td></td>
<td>(0.427)</td>
<td>(0.360)</td>
<td>(0.454)</td>
<td>(0.368)</td>
<td>(0.221)</td>
</tr>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>-0.0909</td>
<td>0.0585</td>
<td>-0.0152</td>
<td>-0.0522</td>
<td>-0.00748</td>
</tr>
<tr>
<td></td>
<td>(0.0975)</td>
<td>(0.0374)</td>
<td>(0.0459)</td>
<td>(0.0638)</td>
<td>(0.0209)</td>
</tr>
<tr>
<td>Burden of Proof Law Index</td>
<td>-0.0670</td>
<td>-0.266</td>
<td>-0.201</td>
<td>-0.260</td>
<td>-0.113</td>
</tr>
<tr>
<td></td>
<td>(0.0661)</td>
<td>(0.0720)</td>
<td>(0.0624)</td>
<td>(0.0852)</td>
<td>(0.0377)</td>
</tr>
</tbody>
</table>

State Fixed Effects  X  X  X  X  X
Year Fixed Effects  X  X  X
States without the Laws Before 1997  X
State-specific Time Trend  X
R-Squared  0.031  0.272  0.420  0.483  0.752
N  1000  1000  1000  460  1000

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of cash benefit payments per covered worker. Standard error are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 1.3: The Effects of Burden of Proof Laws on Medical Benefit Payments

<table>
<thead>
<tr>
<th>Test Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>0.237***</td>
<td>0.110**</td>
<td>-0.00768</td>
<td>-0.0556</td>
<td>0.0195</td>
</tr>
<tr>
<td></td>
<td>(0.0794)</td>
<td>(0.0412)</td>
<td>(0.0398)</td>
<td>(0.0555)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>Burden of Proof Law Index</td>
<td>-0.00843</td>
<td>-0.237***</td>
<td>-0.157***</td>
<td>-0.229**</td>
<td>-0.0821**</td>
</tr>
<tr>
<td></td>
<td>(0.0509)</td>
<td>(0.0620)</td>
<td>(0.0584)</td>
<td>(0.104)</td>
<td>(0.0343)</td>
</tr>
</tbody>
</table>

State Fixed Effects          X           X           X           X           X
Year Fixed Effects            X           X           X           X           X
States without the Laws Before 1997 X
State-specific Time Trend     X
R-Squared                     0.137       0.149       0.342       0.377       0.642
N                              1000        1000        1000        460         1000

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of medical benefit payments per covered worker. Standard error are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 1.4: Decomposition Analysis for the Burden of Proof Index Results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash Benefits Per Covered Worker (log)</td>
<td>5.38</td>
<td>5.01</td>
<td>-0.37</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Burden of Proof Index (Cash Benefits)</td>
<td>0.7</td>
<td>0.9</td>
<td>0.2</td>
<td>-0.201</td>
<td>10.9%</td>
</tr>
<tr>
<td>Medical Benefits Per Covered Worker (log)</td>
<td>5.42</td>
<td>5.13</td>
<td>-0.29</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Burden of Proof Index (Medical Benefits)</td>
<td>0.7</td>
<td>0.9</td>
<td>0.2</td>
<td>-0.157</td>
<td>10.8%</td>
</tr>
</tbody>
</table>

Note: Column (1) and (2) are the means of each variable in 1997 and 2016. Column (3) shows the difference between 2016 and 1997. Column (4) has the regression coefficients for the burden of proof law index in Column (3) of Table 1.2 and 1.3. Column (5) follows Oaxaca and Ransom (1994) and shows the decomposition results on how much of the change in cash benefit payments is explained by the change in each variable.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Benefits Index (log)</td>
<td>1.845***</td>
<td>1.856***</td>
<td>1.875***</td>
<td>1.903***</td>
<td>1.889***</td>
<td>1.749***</td>
</tr>
<tr>
<td></td>
<td>(0.423)</td>
<td>(0.522)</td>
<td>(0.500)</td>
<td>(0.483)</td>
<td>(0.504)</td>
<td>(0.414)</td>
</tr>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>-0.0134</td>
<td>-0.0175</td>
<td>-0.0218</td>
<td>-0.0137</td>
<td>-0.0219</td>
<td>-0.00952</td>
</tr>
<tr>
<td></td>
<td>(0.0467)</td>
<td>(0.0484)</td>
<td>(0.0483)</td>
<td>(0.0451)</td>
<td>(0.0484)</td>
<td>(0.0419)</td>
</tr>
<tr>
<td>Liberal Construction Ban</td>
<td>-0.233*</td>
<td></td>
<td></td>
<td></td>
<td>-0.242**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.119)</td>
</tr>
<tr>
<td>Intoxication Laws</td>
<td></td>
<td>-0.185*</td>
<td>-0.135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.101)</td>
<td></td>
<td></td>
<td></td>
<td>(0.0917)</td>
</tr>
<tr>
<td>Apportionment Laws</td>
<td></td>
<td></td>
<td>-0.195***</td>
<td></td>
<td>-0.119</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0549)</td>
<td></td>
<td>(0.0910)</td>
<td></td>
</tr>
<tr>
<td>Fee Schedule Laws</td>
<td></td>
<td></td>
<td></td>
<td>0.0588</td>
<td>0.0570</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0574)</td>
<td>(0.0563)</td>
<td></td>
</tr>
<tr>
<td>Physician Network Laws</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.144</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.162)</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.396</td>
<td>0.390</td>
<td>0.382</td>
<td>0.376</td>
<td>0.380</td>
<td>0.425</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of cash benefit payments per covered worker. Each regression specification is the standard two-way fixed effects model with state and year fixed effects. Standard error are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 1.6: Decomposition Analysis for Cash Benefit Payments Per Covered Worker

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash Benefits Per Covered Worker (log)</td>
<td>5.38</td>
<td>5.01</td>
<td>-0.37</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Expected Benefits Index (log)</td>
<td>4.34</td>
<td>4.51</td>
<td>0.17</td>
<td>1.749</td>
<td>-80%</td>
</tr>
<tr>
<td>Fatal Injury Rates (log)</td>
<td>1.89</td>
<td>1.52</td>
<td>-0.37</td>
<td>-0.02</td>
<td>-2%</td>
</tr>
<tr>
<td>Liberal Construction Bans</td>
<td>0.14</td>
<td>0.22</td>
<td>0.08</td>
<td>-0.242</td>
<td>5.2%</td>
</tr>
<tr>
<td>Intoxication Laws</td>
<td>0.14</td>
<td>0.2</td>
<td>0.06</td>
<td>-0.135</td>
<td>2.2%</td>
</tr>
<tr>
<td>Apportionment Laws</td>
<td>0.42</td>
<td>0.46</td>
<td>0.04</td>
<td>-0.119</td>
<td>1.3%</td>
</tr>
<tr>
<td>Medical Fee Schedule Laws</td>
<td>0.48</td>
<td>0.72</td>
<td>0.24</td>
<td>0.057</td>
<td>-3.7%</td>
</tr>
<tr>
<td>Physician Network Laws</td>
<td>0.06</td>
<td>0.12</td>
<td>0.06</td>
<td>-0.084</td>
<td>1.4%</td>
</tr>
</tbody>
</table>

Note: Column (1) and (2) are the means of each variable in 1997 and 2016. Column (3) shows the difference between 2016 and 1997. Column (4) has the regression coefficients for each variable in Table 1.5. Column (5) follows Oaxaca and Ransom (1994) and shows the decomposition results on how much of the change in cash benefit payments is explained by the change in each variable. A negative value means the variable change pushed the payments in the opposite direction of the change in the actual payment.
Table 1.7: The Effects of Workers’ Compensation Laws on Medical Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>-0.00869</td>
<td>-0.00844</td>
<td>-0.0120</td>
<td>-0.00322</td>
<td>-0.0157</td>
<td>-0.00276</td>
</tr>
<tr>
<td></td>
<td>(0.0414)</td>
<td>(0.0405)</td>
<td>(0.0415)</td>
<td>(0.0365)</td>
<td>(0.0403)</td>
<td>(0.0346)</td>
</tr>
<tr>
<td>Liberal Construction Ban</td>
<td>-0.0850**</td>
<td>-0.0952**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0423)</td>
<td>(0.035)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intoxication Laws</td>
<td>-0.250**</td>
<td></td>
<td></td>
<td></td>
<td>-0.209***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td></td>
<td></td>
<td></td>
<td>(0.0733)</td>
<td></td>
</tr>
<tr>
<td>Apportionment Laws</td>
<td></td>
<td></td>
<td>-0.0962*</td>
<td>0.0846</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0563)</td>
<td>(0.0702)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fee Schedule Laws</td>
<td></td>
<td></td>
<td></td>
<td>0.0804*</td>
<td>0.0728</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0444)</td>
<td>(0.0437)</td>
<td></td>
</tr>
<tr>
<td>Physician Network Laws</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.271*</td>
<td>-0.238***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.156)</td>
<td>(0.0862)</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.306</td>
<td>0.346</td>
<td>0.305</td>
<td>0.313</td>
<td>0.341</td>
<td>0.387</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of medical benefit payments per covered worker. Each regression specification is the standard two-way fixed effects model with state and year fixed effects. Standard error are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 1.8: Decomposition Analysis for Medical Benefit Payments Per Covered Worker

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Benefits Per Covered Worker (log)</td>
<td>5.42</td>
<td>5.13</td>
<td>-0.29</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Fatal Injury Rates (log)</td>
<td>1.89</td>
<td>1.52</td>
<td>-0.37</td>
<td>0.003</td>
<td>0.4%</td>
</tr>
<tr>
<td>Liberal Construction Bans</td>
<td>0.14</td>
<td>0.22</td>
<td>0.08</td>
<td>-0.095</td>
<td>2.6%</td>
</tr>
<tr>
<td>Intoxication Laws</td>
<td>0.14</td>
<td>0.2</td>
<td>0.06</td>
<td>-0.209</td>
<td>4.3%</td>
</tr>
<tr>
<td>Apportionment Laws</td>
<td>0.42</td>
<td>0.46</td>
<td>0.04</td>
<td>0.085</td>
<td>-1.2%</td>
</tr>
<tr>
<td>Medical Fee Schedule Laws</td>
<td>0.48</td>
<td>0.72</td>
<td>0.24</td>
<td>0.073</td>
<td>-6%</td>
</tr>
<tr>
<td>Physician Network Laws</td>
<td>0.06</td>
<td>0.12</td>
<td>0.06</td>
<td>-0.238</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

Note: Column (1) and (2) are the means of each variable in 1997 and 2016. Column (3) shows the difference between 2016 and 1997. Column (4) has the regression coefficients for each variable in the Column (6) of Table 1.7. Column (5) follows Oaxaca and Ransom [1994] and shows the decomposition results on how much of the change in medical benefit payments is explained by the change in each variable. A negative value means the variable change pushed the payments in the opposite direction of the change in the actual payment.
Chapter 2

Racial, Gender Disparities and State Prosecutors’ Discretion: Evidence from

Blakely v. Washington
I Introduction

Unwarranted disparities are the major controversies in the United States criminal justice system. Research has found, conditional on observable characteristics, Black defendants are more are significantly more likely to be incarceratated, and receive longer sentences, than White counterparts (Rehavi and Starr [2014]). Similarly, men face higher odds of prison time and longer spells in prison than women (Starr [2014]). The causes of these unwarranted disparities are important but less well understood. The disparities might come from unobserved differences across racial and gender groups. Another source of the disparities might be racial and gender discrimination. Broad and largely unreviewable discretion enables judges and prosecutors to treat defendants differently based on their personal tastes. Understanding the causes of unwarranted disparities is critical in crafting effective policy that can ensure equal treatment of defendants regardless of characteristics.

In this paper, we seek to quantitatively estimate the causal effects of restricting prosecutorial discretion on racial and gender disparities in the criminal justice system. The primary challenge in estimating these effects is in measuring changes in prosecutor discretion. Use of discretion is inherently unobservable and may vary drastically in each case. We overcome this hurdle by using a detailed administrative dataset from North Carolina and exploiting an exogenous policy change that raised the state prosecutors’ burdens of proof in successfully seeking sentence enhancements on criminal defendants. This policy change provides the natural experiment we use to measure changes in prosecutor discretion, and sentence enhancements are the key outcome by which we measure the impact of discretion.

Sentence enhancements increase imprisonment periods and are imposed on criminal defendants if they have committed the convicted offense in a severe manner. The severity is evaluated based on legally prescribed aggravating factors, described in detail in Section II of this paper. To impose a sentence enhancement in North Carolina, prosecutors bear the burden of producing sufficient evidence to show the presence of aggravating factors during the crime. The burden of proof in finding aggravating factors was a “preponderance of evidence” standard until the Blakely v. Washington decision abruptly raised the standard to “beyond a reasonable doubt.” In Blakely, the US Supreme Court held that the Sixth Amendment right to a jury trial applies to any fact-finding, except criminal history, that leads to an increase in defendants’ sentence lengths beyond the maximum of the regular sentencing ranges. As a result, if North Carolina prosecutors intend to pursue sentence enhancements, they need to prove the existence of relevant aggravating factors to a jury beyond a
reasonable doubt. This change in the burden of proof effectively restricts prosecutors’ discretion in pursuing sentence enhancements.

Figure 2.2 shows that the Blakely decision causes a sharp and discontinuous drop of around 2 percentage points in the defendants’ probability of receiving sentence enhancements, which translates to a roughly 66% decrease from the pre-Blakely mean. In addition, we find that Blakely did not lead to discontinuities among defendants’ demographics, particularly their criminal history, which supports the validity of a regression discontinuity design to estimate the causal effects of restricting prosecutorial discretion. What’s more, we also show that there is no discontinuous jump in the likelihood of cases being resolved by juries. This suggests that the main mechanism for the discontinuous drop in Figure 2.2 was the prosecutors’ choice of not pursuing sentence enhancements rather than juries applying stricter standards in determining the presence of aggravating factors.

Motivated by this empirical observation, we develop a simple model characterizing how the burden of proof affects prosecutors’ decisions to charge defendants with sentence enhancements. The model assumes that prosecutors seek to maximize expected sentences and predicts that an increase in the prosecutors’ burdens of proof lead to the following two results: (1) an overall decrease in defendants’ probability of being charged with sentence enhancements; (2) a stronger decrease in the probability among subgroups who are discriminated against.

We then employ a regression discontinuity framework to exploit the sharp discontinuity shown in Figure 2.2 and empirically estimate the causal effects of restricting prosecutorial discretion on racial and gender disparities in sentence enhancements. We find strong and robust evidence that Blakely leads to a larger decrease in male defendants’ likelihood of receiving sentence enhancements compared with female counterparts. Prior to Blakely we observe a significant gender gap, with men being 0.7 percentage points, or 28%, more likely to receive sentence enhancements than women after controlling for an extensive set of demographics and fixed effects. Blakely closed the entire male-female gap. In contrast, we do not find evidence suggesting that Black/Hispanic defendants are more likely to receive sentence enhancements than non-Black and non-Hispanic counterparts, pre or post-Blakely. This implies imposing a higher burden of proof has identical effects of defendants from both racial groups.

The estimated null effects of Blakely on racial disparity provide no evidence of racial bias against Black or Hispanic defendants. The changes in gender disparity are primarily driven by two mechanisms: (1) gender differences in the unobserved aggravating factors that affect the take-up of sentence enhancements and (2) prosecutors’ gender bias against men. To clarify the mechanism
channel, we further examine the effects of *Blakely* on gender disparity in sentence enhancements using sub-samples of defendants charged with assault and forgery. Defendants charged with the same offense are likely to receive sentence enhancements for similar reasons, which mitigates the concerns that the gender disparity results are driven by gender differences in unobserved aggravating factors.

We find that *Blakely* had opposite effects on gender disparity in sentence enhancements between the two subgroups of defendants. Among defendants charged with assault, we find a pre-*Blakely* gender gap with men being 71% more likely to receive sentence enhancements. The constraint in prosecutor discretion caused by *Blakely* effectively closes this gap. In comparison, among defendants charged with forgery, we find an opposite pre-*Blakely* gender gap with women being 43% more likely to receive sentence enhancements than men. *Blakely* not only closed the gender gap unfavorable to women but also made men substantially more likely to receive enhancements than women. This sharp comparison suggests that prosecutors tend to associate defendants’ gender with certain types of aggravating factors in pursuing sentence enhancements. The most likely reason why *Blakely* disproportionately benefited men charged with assault is that prosecutors tend to associate men with being cruel or hate crime and decide to pursue sentence enhancements when the evidence is weak. Similarly, prosecutors might associate women with taking advantage of a position of trust and confidence, which explains why *Blakely* disproportionately benefited women charged with forgery.

Our paper makes novel contribution to three different strands of literature in law and economics. There is a rich empirical literature related to racial and gender disparity in the US criminal justice system. Existing studies of racial disparity cover almost all stages of criminal proceeding, including policing [Knowles et al. 2001; Donohue III and Levitt 2001; Gross and Barnes 2002; Anwar and Fang 2006; Fryer Jr 2019; Goncalves and Mello 2020], pre-trial detention [Arnold et al. 2018], prosecution [Tuttle 2019; Sloan 2019], sentencing [Rehavi and Starr 2014] and parole decisions [Anwar and Fang 2015]. Research on gender disparity primarily focus on prosecution and sentencing [Starr 2014; Didwania 2018]. Most relevant to our study is the evidence that prosecutors’ discretionary decisions are the key contributors to the racial and gender sentencing gaps [Rehavi and Starr 2014; Starr 2014]. Our paper rigorously examines defendants’ receipts of sentence enhancements, which is largely driven by prosecutors’ discretion.

Our major contribution comes from the causal evidence of the effects of restricting prosecutorial discretion on racial and gender disparities. The economic literature on expert discretion suggests that giving experts decision-making freedom leads to efficient outcomes and welfare improvement.
But allowing experts to exercise discretion may also lead to undesirable disparities, which in the legal setting, raises concerns over constitutional equal protection. Earlier work also empirically measures prosecutorial discretion in various ways and finds evidence suggesting that restricting prosecutorial discretion leads to changes in prosecutors’ charging behavior and defendants’ outcomes. In this paper, we are able to empirically capture a significant restriction of prosecutorial discretion through a Supreme Court case that requires prosecutors to meet a higher burden of proof standard in pursuing sentence enhancements. This unique setup allows us to apply a regression discontinuity framework to estimate the causal impacts of restricting prosecutorial discretion on gender and racial disparities among a wide range of case outcomes. The analysis sheds new light on prosecutors’ underlying discrimination. In addition, our paper enriches the theoretical literature on prosecutors’ decisions by developing a novel framework that relates prosecutors’ charging decision and underlying bias to burdens of proof.

The rest of the paper is structured as follows. Section II describes the institutional background related to North Carolina sentencing system, Blakely v. Washington, and how Blakely affects the sentencing law practice in North Carolina. Section III reviews the existing literature which this paper contributes to. Section IV presents a simple conceptual model that relates the burden of proof to prosecutors’ decisions to charge defendants with sentence enhancements. Section V describes the data set and the summary statistics. Section VI describes the econometric models and identification strategies. Section VII and VIII describe the main results of the effects of Blakely on gender and racial disparity in sentence enhancements and also the evidence that it is the prosecutors’ behavioral changes that drive the results. Section IX presents the results of the offense-specific effects of Blakely on gender disparity in sentence enhancements using sub-samples of defendants with the same charging offense. Section X concludes with some remarks.

II Institutional Background

II.1 Sentencing in North Carolina

In 1993, North Carolina passed the Structured Sentencing Act as a major legislative effort to address the overpopulation of state prisons and to more effectively combat serious crimes with limited resources. Two years later, the structured sentencing guidelines went into effect,
prescribing the sentencing length of defendants who are convicted of state crimes, both felony and misdemeanor. According to the Structured Sentencing Training and Reference Manual, sentencing determination happens after conviction and requires clear determination of the following elements: the location of the grid box (described below), minimum sentence, and maximum sentence.

Figure 2.1 shows the sentencing guidelines for felony crimes, which was effective from 1995 to 2009. The guidelines consist of “grid boxes” which are determined by two factors: severity of convicted crimes and criminal history. The severity is measured by the assigned offense classes of convicted crimes, which lie on the vertical axis of the guidelines. The state law specifies ten offense classes for felony with class “A” representing the most severe crime and “I” the least severe crime. The horizontal axis measures a defendant’s criminal history points. North Carolina adopted a point system to proxy defendants’ criminal history, where defendants receive certain points based on the severity of prior convictions. For example, defendants are assigned 10 points for a prior conviction of class A felony and 2 points for class H and I felony. Thus, higher criminal history points imply that a defendant had prior convictions of many felonies, a severe felony, or both.

Within each sentencing grid, the guidelines specify three possible ranges for determining defendants’ minimum sentences: presumptive, mitigated, and aggravated ranges. By default, a defendant’s minimum sentence should be within the presumptive range. However, state law prescribes aggravating and mitigating factors which may lead to an upward or downward departure from the presumptive range in determining the defendant’s minimum sentence. The state and the defense respectively bear the burdens of proving the existence of aggravating and mitigating factors. The court chooses the appropriate sentence range by weighing the strength of the present aggravating factors against mitigating factors. After the appropriate range is determined, the judge chooses a specific minimum sentence within the range. Then, the maximum sentence is mechanically determined based on the judge’s choice of the minimum sentence length. The defendant’s final sentence is then determined by judges, who choose from the range of minimum and maximum sentences.

Because enhancements and aggravating factors are not directly observed in our data, we define a sentence as receiving enhancement if the observed minimum sentence comes from the aggravated

---

43The North Carolina sentencing guidelines were amended in 2009. The revised guidelines went into effect from Dec. 1st, 2009 and remain to be effective until today.
44G.S. 15A-1340.14(b)
45G.S.15A-1340.16(d)
46G.S.15A-1340.16(a)
47G.S.15A-1340.16(b)
48G.S. 15A-1340.17(e1)
range. The court’s upward departure from presumptive ranges are conditional upon the State’s valid proof of legally prescribed aggravating factors. However, defendants may admit to certain aggravating factors that have the same legal consequences as those found through formal legal processes. The level of provability varies across the aggravating factors. Some of the aggravating factors are relatively objective and easily provable, such as offenses committed “for the purpose of avoiding or preventing a lawful arrest,” “(involving) a person under the age of 16 in the commission of the crime,” or “(involving) the sale or delivery of a controlled substance to a minor.” In contrast, some of the aggravating factors are quite subjective and difficult to prove, such as an offense being “heinous, atrocious, or cruel,” taking advantage of “a position of trust or confidence...to commit the offense,” or a hate crime enhancement. Changes in the burden of proof matter less for the easily provable aggravating factors. But for the subjective and less provable factors, a higher burden of proof is likely to result in substantial additional costs of searching for extra evidence in order to satisfy the burden.

II.2 Blakely v. Washington

The legal framework of imposing sentence enhancements changed after the Supreme Court case of Blakely v. Washington. In Blakely, the defendant was convicted of second-degree kidnapping through a plea agreement. State prosecutors sought a sentence from the standard range of 49 and 53 months. However, a state court judge unilaterally imposed an exceptional sentence of 90 months, 37 months beyond the maximum of prosecutors’ proposed sentencing range. The judge came to the decision by his own finding of the fact that the defendant committed the crime with “deliberate cruelty,” a statutorily enumerated aggravating factor. The Washington state law allowed judges to sentence defendants above the standard range if “there are substantial and compelling reasons justifying an exceptional sentence.”

The defendant challenged the state judge’s sentencing decision and the Supreme Court sided

49 G.S.15A-1340.16(a1)
50 G.S.15A-1340.16(d)(3),(13)&(16)
51 G.S.15A-1340.16(d)(7),(15)&(17)
52 Under the Washington state law, the second-degree kidnapping is a Class B felony, for which the standard sentencing range falls between 49 and 53 months. See §9.94A.320
53 The “deliberate cruelty” was found based on the following details on how the defendant committed the crime – “He used stealth and surprise and took advantage of the victim’s isolation. He immediately employed physical violence, restrained the victim with tape, and threatened her with injury and death to herself and others. He immediately coerced the victim into providing information by the threatening application of a knife. He violated a subsisting restraining order.” See Blakely v. Washington, 542 U.S. 296 (2004), 301
54 §9.94A.120(2)
with the defendant in a 5-4 decision. Under the majority opinion, the 6th Amendment Right to Jury requires that facts which lead to sentences beyond the standard ranges under the Washington sentencing guidelines are to be found by juries beyond a reasonable doubt. Therefore, the Court struck down the sentencing guidelines in the state of Washington as it unconstitutionally permitted judges to increase defendants’ sentences at their own discretion.

Observers use “revolution” and “earthquake” to describe the disruptive impacts of Blakely on modern sentencing practice (BERMAN 2004; Prescott and Starr 2006). The decision significantly jeopardized the constitutional legitimacy of the mandatory sentencing guidelines at both state and federal levels. As Justice O’Connor said in her dissent opinion, “if the Washington scheme does not comport with the Constitution, it is hard to imagine a guidelines scheme that would.” Similarly, Justice Kennedy predicted that states with similar structured sentencing guidelines to Washington would “scrap everything and start over.”

To comply with Blakely, both state and federal governments introduced significant changes in the legal procedure of determining defendants’ sentences. At the federal level, six months after Blakely, the Court decided United States v. Booker which struck down the federal mandatory sentencing guidelines for violating defendants’ Sixth Amendment Right to jury. The Court held that the federal guidelines remained in place as an “advisory” document, instead of mandatory. At state level, Blakely mainly affected 12 other states with presumptive sentencing structure similar to Washington’s and their responses were quite different (Wool and Stemen 2004). The most common solution was to preserve the integrity of the presumptive sentencing system by adding jury fact-finding requirements into the state laws. Some states like Tennessee switched their sentencing system from presumptive to voluntary where the legally prescribed sentencing ranges were no longer binding. Some states did not make any change.

II.3 The Influence of Blakely on North Carolina

Sentencing in North Carolina follows structured guidelines similar to Washington’s. In both states, judges determine whether to give defendants longer sentences above regular sentencing ranges. However, the judicial discretion in aggravating defendants’ sentences beyond regular ranges is much

---

55 Justice Scalia wrote the majority opinion, which is joined by Justices Thomas, Stevens, Souter and Ginsberg. Justice O’Connor, Kennedy, Breyer, and Rehnquist dissented.
59 According to Stemen and Wilhelm (2005), those states include Alaska, Arizona, California, Colorado, Indiana, Minnesota, New Jersey, New Mexico, North Carolina, Ohio, Oregon, and Tennessee
more constrained in North Carolina. Prior to Blakely, the Washington law gave judges broad discretion to issue “exceptional sentences” based on “substantial and compelling” reasons they perceived. That is why, in the Blakely case, the judge could impose an aggravated sentence despite the prosecutors only seeking a sentence within the standard range. In contrast, pre-Blakely, North Carolina judges were not allowed to impose an sentence enhancement unless state prosecutors requested the enhancement and met the statutory burden of proving the existence of the relevant aggravating factors. More importantly, the presence of aggravating factors were determined by judges, instead of juries.

Blakely intended to limit judicial discretion but incidentally constrained prosecutorial discretion in North Carolina. The decision made it abundantly clear that it was no longer constitutional to have judges determine whether aggravating factors exist in a given case. Rather, juries should make the relevant fact-findings. Therefore, post-Blakely, judges in North Carolina were no longer involved in fact-findings of aggravating factors although they still determine whether to aggravate defendants’ sentences beyond presumptive ranges. Meanwhile, Blakely increased prosecutors’ burdens of proof, from “preponderance of evidence” to “beyond a reasonable doubt,” in proving aggravating factors. Higher burdens of proof and jury requirement inevitably lead to higher costs of pursuing sentence enhancement, such as the additional resources to search for more evidence or to prepare for jury trials. Thus, the Blakely decision effectively constrained prosecutorial discretion in pursuing aggravated sentences.

Blakely was decided in June, 2004. One year later, the state legislature passed a Blakely Bill in June, 2005 and formally amended the laws by requiring juries to determine the presence of legally prescribed aggravated factors by “beyond a reasonable doubt” standard. Defendants may waive their Sixth Amendment right to jury by admitting to the aggravating factors, which are treated as if the aggravating factors were found by juries through the appropriate procedure. However, as law professor Ronald Wright pointed out, the effects of Blakely on North Carolina sentencing practice occurred immediately after the decision. On the one hand, criminal defense attorneys started filing motions requiring juries to determine the presence of aggravating factors if the state sought sentences from aggravated ranges. On the other hand, despite state prosecutors’ preference

60§ 15A-1340.16. Aggravated and Mitigated Sentences.
61The North Carolina state Supreme Court in State v. Allen, 359 N.C. 425 (2005), formally declared the pre-Blakely procedure to determine aggravating factors unconstitutional for violating the defendants’ Sixth Amendment right to jury.
62Under the new amendment, judges are only able to determine the following two aggravating factors: (1) willful violation of probation or parole conditions; (2) prior delinquent adjudication for certain classes of offense. § 15A-1340.16. (12a) & 18(a)
for longer sentences, they started dropping requests for sentence enhancements because the costs of going through a jury trial easily outweighed the benefits. Our empirical analysis exploits this change in prosecutors’ burdens of proof to estimate the causal effects of constraining prosecutorial discretion on racial and gender disparities in sentence enhancements.

III Existing Literature

III.1 Racial and Gender Disparity in the Criminal Justice System

There is a large economic literature which documents racial disparities in various stages of criminal proceedings, including policing (Knowles et al. 2001; Donohue III and Levitt 2001; Gross and Barnes 2002; Anwar and Fang 2006; Gelman et al. 2007; Fryer Jr 2019; Goncalves and Mello 2020), pre-trial detention (Arnold et al. 2018; Arnold et al. 2020), prosecution (Robertson et al. 2019; Sloan 2019; Tuttle 2019), sentencing (Rehavi and Starr 2014) and early release of prisoners (Anwar and Fang 2015; Mechoulan and Sahuguet 2015). In contrast, empirical evidence on gender disparity is relatively thinner and mainly documented in the area of prosecution and sentencing (Mustard 2001; Starr 2014; Didwania 2018).

Prosecutors have enormous discretion in determining how to prosecute crimes. It is policy relevant to understand how prosecutors exercise their discretion and its impacts on defendants' case outcomes. Rehavi and Starr (2014) finds that prosecutors' decisions, particularly charging defendants with mandatory minimum sentences, can explain a significant portion of the racial sentencing gap that can not be explained by observed pre-charge characteristics. However, Rehavi and Starr (2014) does not claim that prosecutors’ charging decisions are driven by racial bias against Black defendants as their decisions could also have been driven by other unobserved factors that are correlated to race. The empirical challenge of studying prosecutors’ racial bias is to control the racial differences in omitted variables. Sloan (2019) exploits as-good-as random assignment of misdemeanor cases to prosecutors to overcome the omitted variables issue because the assignment of cases is uncorrelated with defendants’ race. She documents empirical evidence on prosecutors’ racial bias against defendants of the opposite race that being assigned to prosecutors of the opposite race increases the probability of being convicted of property crimes by 8%. Tuttle (2019) documents prosecutors’ stronger tendency to manipulate the charged amount of crack cocaine drugs against Black or Hispanic defendants in order to charge them with mandatory minimums.

There is another literature on gender disparities. The criminal justice system is a unique setting
where the documented gender disparity suggests unfavorable treatment against male defendants, particularly in prosecution and sentencing. Berdejó (2018) documents systematic gender disparity in prosecution and finds that charges against men are less likely to be dropped or lowered compared to women. Starr (2014) finds prosecutors’ exercise of discretion, such as fact-finding and mandatory minimum sentence charging, is a key contributor to the gender sentencing gap with male defendants receiving longer sentences. In terms of evidence on prosecutors’ gender bias, findings in Didwania (2018) suggest federal prosecutors are likely to charge defendants of the opposite sex with a more severe type of crime.

III.2 Discretion and Prosecutorial Discretion

More generally, this paper relates to the literature examining the desirability of expanding regulatory discretion. Much of the literature suggests that the exercise of discretion leads to efficient outcomes and welfare improvement. Kuziemko (2013) demonstrates that, compared with fixed-sentence regimes, there are significant efficiency gains in having parole boards exercise discretion to determine if a prisoner should be released earlier. Parole officers appear to be well-trained to identify and grant early release to low-risk offenders who are less likely to recidivate than those released after serving a fixed amount of sentences. Similarly, in the context of punishing water polluters, Kang and Silveira (2018) finds that regulators can flexibly adapt to local preferences when exercising discretion. Their counterfactual analysis suggests removing such enforcement discretion leads to a significant welfare loss.

However, expanding regulatory discretion might amplify the existent bias underlying regulators’ behavior. Regulators’ decision bias is particularly concerning in the criminal justice system. A more accurate understanding of the relationship between undesirable disparities and exercise of discretion by judges or prosecutors can better serve the public interests (Engen, 2009). A few papers study the causal relationship between judicial discretion and racial disparity by using Booker v. United States as a natural experiment (Yang 2015, Starr and Rehavi 2013). Booker held that federal judges were no longer required to follow sentencing guidelines in determining defendants’ sentences. Empirical work has shown that such expansion of judicial discretion increases inter-judge variation in sentencing lengths (Scott 2010, Yang 2014) and also exacerbates the racial sentencing gap by increasing Black defendants’ sentencing lengths by about 2 months (Yang 2015).

There is another set of economic literature examining prosecutorial discretion, both theoretically and empirically. The theoretical literature develops various models on how prosecutors make deci-
sions by optimizing some objective function, such as maximizing expected sentences (Landes 1971) or maximizing social welfare (Grossman and Katz 1983; Reinganum 1988). The challenge in the empirical literature is to quantify prosecutorial discretion. Yang (2016) overcomes the challenge by examining how prosecutors respond to vacancies of federal benches. The results suggest that, due to this restriction, prosecutors are more likely to dismiss filed charges and to decline to prosecute by offering defendants favorable plea deals that do not involve imprisonment. There is also a set of papers measuring the effects of criminal procedure and policy changes on prosecutors’ discretionary charging decisions (Bjerk 2005; Kessler and Piehl 1998; Prescott 2006a; Prescott 2006b). Prescott (2006a) finds that prosecutors reduce the number of charging counts by around 10% in response to Apprendi v. New Jersey, a Supreme Court’s decision strengthening the constitutional protection of defendants’ right to a jury trial.

III.3 Evaluation of Changes in Legal Standards

Our major identifying variation comes from an increase in the prosecutor’s burden of proof to pursue sentence enhancements. The Blakely decision requires prosecutors to produce more facts in order to establish aggravating factors. Our analysis also contributes to a small law and economics literature on empirical evaluation of changes in legal standards.

Priest and Klein (1984) developed a seminal economic model which describes how litigants with divergent expectations make optimal decisions to litigate given disputes. One important result of the model is the “Priest & Klein hypothesis,” which predicts that plaintiffs’ win rate stays at 50% regardless of the applicable legal decision standard. The rationale is that litigated disputes are self-selected based on the decision standard. If a change in legal standard makes it tougher for plaintiffs to win, we might expect plaintiffs’ win rate to fall. However, the tougher legal standard also makes plaintiffs to litigate disputes of higher quality, which increases plaintiffs’ probabilities to win. Overall, the plaintiffs’ equilibrium win rate stays at 50%.

The Priest & Klein hypothesis has empirical support in the empirical literature on how litigation outcomes are affected by changes in pleading standards caused by two Supreme Court cases, Bell

---

63 There is empirical evidence supporting the validity of the objective of maximizing expected sentences. For example, prosecutors seeking longer sentences tend to have more favorable career outcomes after leaving prosecutors’ office (Boylan 2005).

64 Yang argued that judicial vacancies restrict prosecutorial discretion by increasing the costs of prosecution. With fewer federal judges, criminal cases are likely to be resolved more slowly, which increases the probability of speedy trial violation and other adverse consequences, such as leading evidence becoming stale and loss of memories of key witnesses.

65 530 US 466 (2000). The Court held that facts which increases defendants’ sentence length beyond statutory maximum should be proved to jury beyond a reasonable doubt.
Atlantic Corp v. Twombly and Ashcroft v. Iqbal (Hubbard 2013; Hubbard 2017; Gelbach 2011; Gelbach 2016). Twombly and Iqbal instituted a “plausibility” pleading standard, which requires more factual pleading from plaintiffs in order to survive defendants’ motions to dismiss. The empirical analysis documents no evidence that a more stringent factual pleading standard leads to significant changes in the likelihood of courts’ grant of motion to dismiss (Hubbard 2013) and motion for summary judgment (Gelbach 2016).

In contrast to the Priest & Klein hypothesis, our results show that changing the burden of proof standard leads to a sharp reduction in defendants’ probabilities of receiving sentence enhancements, as Figure 2.2 clearly shows. This result appears to contradict the Priest & Klein hypothesis as a change in decision standard significantly changes plaintiffs’ “win rate,” measured by not receiving sentence enhancements. The most likely explanation is that the change in prosecutors’ burdens of proof did not lead to significant sample selection bias due to police’s decisions to arrest and prosecutors’ initial charging decisions.

IV A Simple Model of Prosecutors’ Decision Making

This section presents a simple conceptual model on prosecutors’ decisions to charge sentence enhancements. The model yields empirically testable predictions of prosecutor response to a higher burden of proof.

IV.1 Main Model

The framework models the decision of the prosecutor to pursue sentence enhancement. Specifically, prosecutors decide if they will charge defendant $i$ with a sentence enhancement $S$, a random variable with probability density function $f(s)$ and cumulative distribution function $F(s)$. $\theta_i(\pi_i, B)$ is the probability that case $i$ merits a sentence enhancement. We model the probability as a function of the following two elements: (1) specific characteristics of case $i$, both observed and unobserved, represented by $\pi_i$, and (2) burden of proof that a prosecutor needs to meet, $B$. $B$ does not have subscript $i$ as a burden of proof standard, either “preponderance of evidence” or “beyond a reasonable doubt,” applies to all cases regardless of the characteristics. In addition, we impose a monotonicity assumption that $\theta_i(\pi_i, B)$ is decreasing in $B$, meaning that as the burden of proof increase.

---

550 U.S. 544 (2007)
556 U.S. 662 (2009)
Similarly, Yuan and Fishback (2020) also finds that increasing workers’ burdens of proof restricts their access to workers’ compensation benefits measured by cash and medical benefits payment per covered worker.
increases, the probability of meeting the burden decreases. For simplicity, assuming that $B$ is defined on a continuous range and $\theta_i(\pi_i, B)$ is differentiable in $B$, the monotonicity assumption can be written as the following partial derivative, $\frac{\partial \theta_i(\pi_i, B)}{\partial B} < 0$.

$C_i(\pi_i, B)$ is the prosecutor’s cost of charging individual $i$ with a sentence enhancement. This is also a function of case characteristics, $\pi_i$, and the burden of proof, $B$. We impose a similar monotonicity assumption that $C_i(\pi_i, B)$ is increasing in $B$. This implies that an increase in the burden of proof leads to a higher cost of charging sentence enhancements, which refers to the additional resources and effort of finding evidence needed to meet the higher burden. For simplicity, the monotonicity assumption can be written as the following partial derivative, $\frac{\partial C_i(\pi_i, B)}{\partial B} > 0$.

Prosecutors’ expected payoff of charging $S$ can be written as $\theta_i(\pi_i, B)S - C_i(\pi_i, B)$, the difference between the expected benefit, $\theta_i(\pi_i, B)S$, and the cost, $C_i(\pi_i, B)$. We follow the general assumption in the literature that prosecutors’ objective is to maximize expected sentencing length (Landes 1971; Grossman and Katz 1983). Therefore, in deciding whether or not to charge defendant $i$ with a sentence enhancement $S$, prosecutors solve the following decision problem:

$$\max \{ \theta_i(\pi_i, B)S - C_i(\pi_i, B), 0 \}$$

We can then write the probability of defendant $i$ being charged with sentence enhancement $S$ as follows:

$$P_i(\pi_i, B) = Pr(\theta_i(\pi_i, B)S > C_i(\pi_i, B)) = Pr \left( S > \frac{C_i(\pi_i, B)}{\theta_i(\pi_i, B)} \right) = 1 - F \left( \frac{C_i(\pi_i, B)}{\theta_i(\pi_i, B)} \right)$$

The model implies that a prosecutor’s decision to charge individual $i$ with sentence $S$ follows a cut-off rule: if $S$ is larger than the cut-off $\frac{C_i(\pi_i, B)}{\theta_i(\pi_i, B)}$, the sentence enhancement will be pursued.

**Proposition 1.** If there is an increase in $B$, all else equal, prosecutors are less likely to charge defendants with sentence enhancements.

The proof involves taking the partial derivative of $P_i(\pi_i, B)$ with respective to $B$. It is straight-
forward to show that the cut-off, \( \frac{C_i(\pi, B)}{\theta_i(\pi, B)} \), is increasing in \( B \). An increase in \( B \) results in higher cut-off, which lowers the probability, \( P_i(\pi, B) \).

The model also makes it clear that a higher burden of proof affects prosecutors’ choice in pursuing sentence enhancements through two main channels. On the one hand, it reduces expected benefits by decreasing the probability of success. On the other hand, it increases the overall costs of pursuing sentence enhancements because meeting a higher burden of proof requires additional efforts and resources to search for more evidence.

IV.2 Modeling Prosecutors’ Bias

We expand this simple framework to incorporate prosecutors’ bias against a subgroup of defendants. A common economic method implements agents’ bias with group-specific decision cut-offs [Becker 2010; Knowles et al. 2001; Anwar and Fang 2006]. To simply the notation, we let \( G(\pi, B) = \frac{C_i(\pi, B)}{\theta_i(\pi, B)} \) represent the equilibrium cut-off of prosecutors’ decisions to charge sentence enhancements. \( G_A(\pi, B) \) and \( G_{A'}(\pi, B) \) are prosecutors’ decision cut-off against defendants in subgroup \( A \) and \( A' \). Then, we use the following condition to describe prosecutors’ gender bias against defendants of subgroup \( A \):

\[
\forall i, \ G_A(\pi, B) < G_{A'}(\pi, B)
\]

Conditional on the same case characteristics and burdens of proof, prosecutors’ bias against defendants in subgroup \( A \) is reflected by a strictly lower cut-off in charging defendants of subgroup \( A \) with sentence enhancements. Prosecutors’ bias might come from different sources. Prosecutors’ lower cut-off might come from their distaste against defendant in subgroup \( A \), which is called taste-based discrimination. Statistical discrimination is also plausible. Prosecutors’ prior experience might inform them that, compared with defendants in subgroup \( A' \), those in subgroup \( A \) are more likely to merit a sentence enhancement, which results in the lower cutoff. Prosecutors’ bias might also be a combination of both taste-based and statistical discrimination. This paper does not attempt to distinguish between taste-based and statistical discrimination.

We are primarily interested in expressing the differential effects of a higher burden of proof on
prosecutors charging decisions against defendants in two different subgroups. Suppose there is an increase in the burden of proof from $B^L$ to $B^H$, such that $B^H > B^L$. The effect of a higher burden on the probability of charging defendants in subgroup $A$ with a sentence enhancement:

$$P_i(B^H|\pi_i, A) - P_i(B^L|\pi_i, A) = F(G_A(\pi_i, B^L)) - F(G_A(\pi_i, B^H))$$

To draw a more clear comparison between defendants in the two subgroups, we further rewrite that $G_{A'}(\pi_i, B) = G_A(\pi_i, B) + \Delta C(\pi_i)$ and $\Delta C(\pi_i) > 0$. The effects of a higher burden of proof on the likelihood of defendants in subgroup $A'$ being charged with a sentence enhancement can be written as follows:

$$P_i(B^H|\pi_i, A') - P_i(B^L|\pi_i, A') = F(G_{A'}(\pi_i, B^L)) - F(G_{A'}(\pi_i, B^H))$$

$$= F(G_A(\pi_i, B^L)) - F(G_A(\pi_i, B^H))$$

$$+ \int_{G_A(\pi_i, B^L)}^{G_A(\pi_i, B^L) + \Delta C(\pi_i)} f(s)ds - \int_{G_A(\pi_i, B^H)}^{G_A(\pi_i, B^H) + \Delta C(\pi_i)} f(s)ds$$

Compared with defendants in subgroup $A'$, the differential effects of a higher burden of proof on the likelihood of defendants in subgroup $A$ being charged with a sentence enhancement can be written as follows:

$$[P_i(B^H|\pi_i, A) - P_i(B^L|\pi_i, A)] - [P_i(B^H|\pi_i, A') - P_i(B^L|\pi_i, A')] =$$

$$\int_{G_A(\pi_i, B^L)}^{G_A(\pi_i, B^L) + \Delta C(\pi_i)} f(s)ds - \int_{G_A(\pi_i, B^H)}^{G_A(\pi_i, B^H) + \Delta C(\pi_i)} f(s)ds$$

Using the assumption that $f(S)$ is decreasing in $S$ and an earlier result that the optimal cut-off $g_i(\pi_i', B)$ is increasing in $B$, we can show that $[P_i(B^H|\pi_i, A) - P_i(B^L|\pi_i, A)] - [P_i(B^H|\pi_i, A') - P_i(B^L|\pi_i, A')] < 0$. We can apply this set of analysis to model prosecutors’ gender bias against male defendants and racial bias against Black/Hispanic defendants. Then, we state the following two propositions.

**Proposition 2.** If gender bias against male defendants is present, an increase in the burden of proof decreases the probability of being charged with a sentence enhancement more for male


**Proposition 3.** If racial bias against Black or Hispanic defendants is present, an increase in the burden of proof decreases the probability of being charged with a sentence enhancement more for Black or Hispanic defendants than for non-Black and non-Hispanic defendants.

*Blakely v. Washington* introduces a plausibly exogenous change in prosecutors’ burden of proof in pursuing a sentence enhancement. Exploiting this variation, we empirically test the theoretical predictions developed here through a regression discontinuity design shown in section VII. It is also worth highlighting that it is hard to empirically test Proposition 2 & 3 because it requires researchers to control for both observable and unobservable characteristics. Suppose $\pi_i^*$ represents the observable characteristics of case $i$. Empirically, we are only able to identify $[P_i(B^H|\pi_i^*, A) - P_i(B^L|\pi_i^*, A)] - [P_i(B^H|\pi_i^*, A') - P_i(B^L|\pi_i^*, A')]$. Even if we are able to empirically demonstrate that

$$[P_i(B^H|\pi_i^*, A) - P_i(B^L|\pi_i^*, A)] - [P_i(B^H|\pi_i^*, A') - P_i(B^L|\pi_i^*, A')] < 0$$

We are reluctant to claim that the result is due to prosecutors’ bias against subgroup $A$ because there might be group-specific unobserved characteristics that can explain this result. We are unable to exhaustively control for group-specific unobservable characteristics in our empirical analysis.

V Data and Summary Statistics

Our main data source consists of criminal case records from the Administrative Office of Courts (AOC) in North Carolina. The dataset includes all of the post-arrest charges filed in the state courts. The data has a rich set of individual characteristics including names, date of birth, sex, race, criminal history, attorney representation, and the type of crimes being charged. We also observe outcomes at various stages of each case, such as prosecutors’ dismissal of initial charges, plea bargain, conviction, minimum and maximum sentences. We also supplement the main data source with county-level arrest data from the Uniform Crime Report (UCR) and unemployment data from the Bureau of Labor Statistics (BLS). After extensive data cleaning, we restrict our analysis sample to all of the criminal “cases” charged with felony offenses from 2001 to 2007, three

---

71We do not observe actual sentence received and served.
Table 2.1 shows the summary statistics of our analysis sample, which consists of 213,264 cases. 84% of the defendants are male. 58% of the defendants are Black or Hispanic and 40% are white. The average age is around 30 years old. 24% of the individuals are represented by private attorneys and 22% represented by public defenders. The mean of criminal history points is 4.6. As to main case outcomes, 83% of the cases are resolved through plea bargains, and the rate increases to 95.9% if we exclude the set of cases voluntarily dismissed by prosecutors before plea bargain. Defendants received incarceration sentences in 32% of the cases. The average minimum sentence length is around 20 months. The average county-level arrest rate is 536 per 100,000 county population. The average county-level unemployment rate is 5.8.

The dataset does not directly include information on sentence enhancements and findings of aggravating and mitigating factors. However, we develop a measure of sentence enhancements by implementing the following two steps: (1) using the criminal history points and the class of convicted offenses to identify the relevant sentencing grid shown in Figure 2.1 and (2) matching the observed minimum sentence lengths to the applicable sentencing ranges to determine if there is an upward or downward deviation from the presumptive sentencing range. We define that a defendant receives a sentence enhancement if the observed minimum sentence length comes from an aggravated range. We find that defendants received sentence enhancements in 2% of the cases due to the finding of aggravating factors. Minimum sentences come from mitigated range in 8% of the total cases, respectively.

Figure 2.2 plots the monthly averages of cases where defendants receive sentence enhancements from 2001 to 2007, which are calculated using the raw data. The figure shows that Blakely leads to a sharp and discontinuous drop in defendants’ likelihood of receiving sentence enhancements. The drop is unlikely to be driven by underlying trends or other factors. Figure 2.B.5 shows the monthly averages of “residualized” sentence enhancements. The discontinuity at the Blakely cutoff remains sharp. Table 2.C.1 shows the summary statistics before and after Blakely. There does not appear to be any significant changes in defendants’ characteristics after the Blakely decision.

---

72 We identify “case” by using pairs of individual identifier and case disposition date. Please see Appendix A for more details on data cleaning and “case” definition.
73 Unemployment rate represents the number of unemployed people as a percentage of the labor force (the sum of employed and unemployed).
74 In calculating the monthly averages, we use the total number of cases charged with felony offenses in a given month as the denominator and the total number of cases with sentence enhancements imposed as the numerator.
75 The “residualized” sentence enhancements are the residuals from the regression model which regresses a dummy of sentence enhancements on defendants’ observable characteristics, guidelines fixed effects, and county fixed effects.
Figure 2.3 shows the gender-specific monthly averages of sentence enhancements from 2001 to 2007. The graph illustrates several important points for our analysis. First, it shows a clear gender gap with men being more likely to receive sentence enhancements than women in both pre and post Blakely periods. Second, the gender gap becomes significantly smaller in post-Blakely periods. Third, the month-to-month variation appears to be larger in pre-Blakely periods, suggesting that a proof by beyond a reasonable doubt also improves cohesiveness in defendants’ sentence enhancements outcomes. Figure 2.B.6 shows the gender gap in “residualized” sentence enhancements. Conditional on observed characteristics, men are still more likely to receive sentence enhancements during the pre-Blakely periods. But, the gap no longer exists in post-Blakely periods, which implies that the post-Blakely gender gap in Figure 2.3 can be entirely explained by observable characteristics.

Figure 2.4 shows the race-specific monthly shares of cases with sentence enhancements from 2001 to 2007 comparing Black or Hispanic defendants to non-Black or non-Hispanic defendants. In both pre and post Blakely periods, there are no clear racial disparities between these groups. Thus, Blakely appears to have similar effects for defendants of different races. The race-specific averages of “residualized” sentence enhancements in Figure 2.B.7 convey similar information.

VI Empirical Strategy: Regression Discontinuity Design

We estimate the following regression discontinuity model:

\[ Y_{ijkt} = \beta_0 + \beta_1 Z_i + \beta_2 1[Blakely] + \beta_3 (Z_i \times 1[Blakely]) + \beta_4 \tau_t + X_i \gamma + Grid_{jk} + \theta_c + \epsilon_{ijkt} \]

\(Y_{ijkt}\) is the main outcome variable, indicating whether the minimum sentence is aggravated for individual \(i\) with criminal history level \(j\), convicted to offense level \(k\) in county \(c\), at time \(t\). Defendant’s take-up of sentence enhancement is the judge’s decision, but it is conditional upon the prosecutor’s valid proof of aggravating factors. Section VIII discusses the reasons why Blakely-induced changes in defendants’ take-up of sentence enhancements is mainly driven by the prosecutor’s decision to not pursue them and not a result of changes in judicial discretion.

\(Z_i\) is either the race or the gender of defendant \(i\). \(1[Blakely]\) is a binary variable equal to 1 in all time periods after June 2004 when Blakely was decided and 0 otherwise, which constitutes the “discontinuity” for our identification. \(\tau_t\) is a flexible function of month, the running variable {Boomhower} 2019. \(X_i\) is a vector of observed control variables, including defendants’ demographi-
ics and county-level characteristics. Defendants’ demographics include age, criminal history points, types of attorney representation, and charged offenses. County-level characteristics include population, arrest rate, and unemployment rate. We also create a vector of guidelines fixed effects, $\text{Grid}_{jk}$, by interacting criminal history class $j$ with convicted offense $k$, which controls unobserved characteristics of defendants sentenced at different sentencing guidelines grids in Figure 2.1. $\theta_c$ represents the county fixed effects, controlling for geographical differences across counties. $\epsilon_{ijkct}$ is an error term. The standard errors are clustered at the month level.

If $Z_i$ indicates whether defendant $i$ is male\(^{76}\), then $\beta_1$ estimates the pre-\textit{Blakely} gender disparity in receiving sentence enhancements. $\beta_2$ measures the effects of \textit{Blakely} on sentence enhancements that are common to both male and female defendants. $\beta_3$ measures the differential effects of \textit{Blakely} on male defendants. $|\frac{\beta_3}{\beta_1}|$ measures of the percentage change in gender gap in sentence enhancements after \textit{Blakely}.

According to our theoretical model of prosecutor behavior, the estimate of $\beta_2$ tests Proposition 1 and is expected to be negative as \textit{Blakely} empirically captures an exogenous increase in prosecutors’ burdens of proof. Because $\beta_3$ captures the differential effect of \textit{Blakely} on subgroups of defendants, it is related to the hypothesis made in Proposition 2 & 3. Take gender as an example. Our model suggests that a negative estimate of $\beta_3$ implies prosecutors’ gender bias against men only if men and women are the same in unobserved characteristics, an assumption unlikely to hold empirically. Alternatively, the negative sign of $\beta_3$ could be driven by gender differences in the unobserved aggravating factors. In section IX, we separately examine the effects of \textit{Blakely} using sub-samples of defendants who are charged with the same felony offenses and who are more likely to receive enhancements based on similar aggravating factors. The offense-specific results can shed more light on potential gender bias from prosecutors.

Our empirical strategy is to exploit the timing of the \textit{Blakely} decision through a regression discontinuity design. Following \textit{Lee and Lemieux} (2010), we estimate the key parameters using local linear regressions and also perform flexible polynomial regressions as robustness check. In order to interpret $\beta_2$ and $\beta_3$ as causal estimates, we need to make the following assumptions: (1) conditional on controls and fixed effects, the timing of \textit{Blakely} is exogenous: $\text{cov}(1[\text{Blakely}], \epsilon_{ijkcm}|X_i, \text{Grid}_{jk}, \theta_c, \lambda_m) = 0$; (2) gender and racial groups have parallel pre-trends prior to the \textit{Blakely} decision, parallel trends assumption.

\(^{76}\)The key parameters can be interpreted to explain racial disparity in the same way if $Z_i$ indicates whether individual $i$’s race is Black or Hispanic.
The main threat to identification comes in the form of potential manipulation. To implement a valid regression discontinuity design, there needs to be no systematic manipulation at the cut-off (Lee and Lemieux 2010). In our context, the identification could be undermined if criminal justice agents strategically respond to Blakely, which would lead to systematic differences between the pre-Blakely and post-Blakely samples. We argue that such selection issues are mitigated in this setting.

There are two possible sources of selection bias issues related to criminal defendants. First, potential criminals may strategically adjust their criminal behavior in response to Blakely as they are less likely to be punished with enhancements. This is highly unlikely because defendants would need to have fairly technical knowledge of aggravating factors, prosecutors' burdens of proof in pursuing sentence enhancements, and being aware of the timing of Blakely. Second, the defendants' attorney might strategically manipulate the case disposition dates so that the new law could favorably apply to their clients. This concern is partially mitigated because Blakely was a 5-4 narrow decision and oppose to the expectation from legal professionals who anticipated the Court not to decide the case in a way that might radically transform the modern sentencing practice (Berman 2004; Pfaff 2006).

Additionally, both prosecutors and law enforcement may respond to Blakely in a way that undermines our identification strategy. For example, prosecutors might respond to Blakely by being less likely to lower defendants’ felony charges to misdemeanors in order to “sufficiently” punish defendants. Such behavioral change might cause sample attrition bias because sentence enhancements do not exist for defendants convicted of misdemeanors in North Carolina. Also, prosecutors might charge defendants more aggressively in order to circumvent the new procedural barrier in imposing sentence enhancements (Bibas 2001; King and Klein 2001). This aggressive charging behavior might change the post-Blakely case composition within the guidelines grid cells. Finally, in response to Blakely, prosecutors might collude with law enforcement to prosecute different types of criminal cases, which might also cause sample selection bias.

We implement a number of checks to make sure that potential manipulation does not undermine our identification. First, we run a McCrary test at the cutoff date (McCrary 2008). The results in Figure 2.5 suggest that there are no significant changes in the density of case volumes at the cut-off date, which provides no evidence of manipulation. In addition, we empirically test whether Blakely leads to discontinuities in defendants’ observable characteristics. These results are shown in Figure 2.B.1. We find no discontinuities in a wide range of defendants’ demographics, particularly criminal history points, at the Blakely cutoff date. This suggests that the types of defendants being
charged are essentially the same before and after *Blakely*. Furthermore, to address the concerns over sample attrition bias we examine the pre- and post-*Blakely* differences in defendants’ likelihood of pleading guilty to misdemeanors. The results in Figure 2.B.4 provide no evidence that there is any significant change. Also, Figure 2.B.3 shows that *Blakely* does not lead to significant changes in the county-level arrest rate, which suggests the police did not systematically change their behavior in response to *Blakely*. Finally, post-*Blakely* changes in case composition within guidelines grid cells are unlikely to undermine our identification. Our identification does not reply on making pre and post *Blakely* comparison within guidelines grids. The results are robust across various specifications regardless of controlling guidelines fixed effects.

Another threat is that the differential impacts across racial and gender groups are driven by underlying trends, instead of the policy change. To formally test whether the changes in disparities are driven by preexisting trends, we estimate the following event-study specification:

\[
Y_{ijkct} = \beta_0 + \beta_1 Z_i + \beta_2 1[\text{Blakely}] + \sum_t \eta_t (Z_i \times 1[\text{Blakely}]_t) + X_i \gamma + Grid_{jk} + \theta_c + \epsilon_{ijkct}
\]

In the above event-study model, 1[Blakely]_t means t period(s) away from the timing of the *Blakely* decision, June of 2004. If t is negative, 1[Blakely]_t denotes |t| periods before the decision, and vice versa. The time period is measured at the biannual level. Therefore, \(\eta_t\) is a vector of period-specific coefficients which measure how gender/racial disparities change over time due to the *Blakely* decision. The estimates of \(\eta_t\) further assess the validity of the parallel pre-trends assumption for identifying \(\beta_3\) in the main regression discontinuity specification.

**VII Main Results**

**VII.1 The Overall Effects of *Blakely***

Table 2.3 shows the results of how raising prosecutors’ burden of proof affects defendants’ likelihood of receiving sentence enhancements. In order to show the robustness of the results, we separately estimate the following five different regression specifications: the specification in Column (1) only includes a *Blakely* dummy, Column (2) adds a rich set of control variables at both individual and county levels, the Column (3) & (4) specification adds county and month-of-year fixed effects, and Column (5) further adds guidelines grid fixed effects. We perform local linear regression analysis.
in all of the specifications except Column (4) where we estimate a flexible polynomial regression model. The preferred specification is Column (3) which includes rich sets of controls and does not require making pre and post Blakely comparison within guidelines grid cells.

The regression results affirm the graphical evidence in Figure 2.2 which shows a discontinuous drop in sentence enhancements. The point estimate in Column (3) is -1.9 percentage points and is statistically significant at the 1% level. The pre-Blakely mean of sentence enhancements is 0.029. The result is consistent with the theoretical prediction in Proposition 1 and implies that Blakely decreases the defendants’ likelihood of receiving sentence enhancements by 65.5%, a large effect. In order words, Blakely reveals that 65.5% of the pre-period sentence enhancements were based on weak proof of aggravating factors which were unlikely to survive proof by beyond a reasonable doubt. The estimate in Column (4) suggests that the result is robust to different functional form of the running variable. Compared with Column (3), Column (5) specification further add guidelines fixed effects and produces the same point estimate, which suggests that post-Blakely changes in case composition within guidelines grid cells do not affect the results. Table 2.C.4 shows that the results are also robust to different bandwidths choices. Using a 6-month bandwidth produces a slightly larger estimate than other bandwidth choices.

Although we do not observe the specific aggravating factors associated with each sentence enhancement, it is reasonable to speculate that these large effects are mainly driven by the enhancements based on aggravating factors that are relatively subjective and not easily provable. For example, proving a hate crime is difficult and requires extensive factual details on defendants' motivation beyond victims' race, religion, or nationality. Blakely raised the evidentiary threshold of establishing hate crime even higher. The costly process of searching for additional evidence to meet the higher standard is likely to deter prosecutors from pursuing hate crime enhancements. This idea of marginal cases being those with weaker evidence is supported by the analysis in Section IX.

VII.2 The Effects of Blakely on Gender Disparity in Sentence Enhancements

Figure 2.3 presents graphical evidence on how Blakely affects gender disparity in sentence enhancements. Pre-Blakely there is a large gender gap with men being more likely to receive sentence enhancements than women. The Blakely decision closed around half of the pre-existing gender

77 In the polynomial regression model, we add second-degree polynomials in the running variable, month, on each side of the Blakely decision date.
gap. Figure 2.B.6 shows that more striking evidence that, conditional on observed characteristics, the gender gap reversed after Blakely and women became slightly more likely to receive enhancements than men. In addition, the figures also show that the month-to-month variation in sentence enhancement rates is larger pre-Blakely when prosecutors had more discretion to pursue sentence enhancements, which is consistent with prior studies demonstrating that allowing more discretion leads to larger discrepancies in defendants outcomes (Scott 2010, Yang 2014).

Table 2.4 presents the results for the effects of Blakely on gender disparity in sentence enhancements. The coefficient for the male dummy variable is the pre-Blakely gender gap in the likelihood of receiving a sentence enhancement. All of the point estimates are positive and statistically significant at 1% level, suggesting that male defendants have a higher probability of receiving sentence enhancements compared to their female counterparts. This result is consistent with findings in the literature on gender disparity against male defendants in criminal case outcomes (Mustard 2001, Starr 2014). Using the estimate of 0.0124 in Column (1) as the baseline gap, the estimates in Column (2) and (5) imply the individual and county-level demographics and guidelines fixed effects separately account for 37% and 6% of the gap. County fixed effects and month-of-year fixed effects explain very little of the gender gap. Together, observable characteristics explain 43% of the observed gender gap. The remaining pre-Blakely gender gap of 0.71 percentage points in Column (5) suggests that, conditional on observed characteristics, men are 28% more likely to receive enhancements than women. This 28% probability disparity is driven either by omitted variables or by prosecutorial and/or judicial discretion.

Based on the estimates in Column (5), the coefficient for the 1[Blakely] dummy is -0.012 and it is statistically significant at the 1% level. This coefficient suggests that Blakely leads to a 1.2 percentage point decrease in the likelihood of receiving sentence enhancements for both male and female defendants. The coefficient for the interaction term (1[Blakely] * Male) implies an additional 0.81 percentage point decrease in male defendants’ likelihood of receiving sentence enhancements. This result reveals that, pre-Blakely, prosecutors have a stronger tendency to go after male defendants for sentence enhancements based on weak evidence that is unlikely to survive proof by beyond a reasonable doubt.

The ratio of the two estimates for 1[Blakely] * Male and Male is around -1.14, which suggests that Blakely closed the entire pre-existing unexplained gender gap and made women slightly more likely to receive enhancements than men during the post-Blakely periods. This implies the entire pre-existing gender disparity was attributed to allowing judicial discretion in finding aggravating
factors and a lack of constraints on prosecutorial discretion to pursue sentence enhancements under the previous “preponderance of evidence” standard. The results are robust across each specification and also different bandwidth choices according to Table 2.C.4.

Although this result is consistent with the theoretical prediction in Proposition 2, we are reluctant to claim it as evidence of gender discrimination. Alternatively, the results may be driven by gender-differences in the unobserved reasons related to the take-up of sentence enhancements. For example, it is plausible that men are more likely to receive sentence enhancements for committing a hate crime or being heinous and cruel in their criminal activities. Men could have disproportionately benefited from *Blakely* because it is inherently difficult to prove a hate crime or cruelty by a beyond a reasonable doubt standard, which is unrelated to prosecutors’ discrimination against men.

Figure 2.6 shows the event study plots on how gender disparity in sentence enhancements changes over time. The time period is measured at the semi-year level and time 0 indicates the second half of 2004. The figure suggests the post-*Blakely* change in gender disparity is not driven by pre-existing trends. All of the estimated coefficients for pre-*Blakely* periods are statistically indistinguishable from 0. In contrast, the coefficients for all of the post-*Blakely* periods are negative and statistically significant at 1% level, suggesting that the effects of *Blakely* on gender disparity are immediate and persistent over time.

**VII.3 The Effects of *Blakely* on Racial Disparity in Sentence Enhancements**

Figure 2.4 shows the race-specific monthly share of defendants who receive sentence enhancements. There is not any graphical evidence of a racial disparity in sentence enhancements in both pre and post *Blakely* periods. Thus, *Blakely* appears to have identical effects on Black/Hispanic defendants and all other defendants. These observations are largely corroborated in Figure 2.B.7 which implies that, conditional on a rich set of observed characteristics, there is not a clear racial gap in the residualized sentence enhancements, both pre and post *Blakely*.

Table 2.5 presents the OLS results for the effects of raising prosecutors’ burdens of proof on the racial disparity of sentence enhancement. The estimates are similar across each specification. The point estimates for the Black/Hispanic dummy indicate very little evidence of a racial disparity in receiving sentence enhancements, which is consistent with the graphical evidence in Figure 2.4.

---

78 In Table 2.C.4, the results across different bandwidth choices consistently show that *Blakely* closed the entire pre-existing gender gap in sentence enhancements.
All of the point estimates are of small magnitude and statistically indistinguishable from 0. These results draw a contrast to our expectations and the prior findings in the literature. Rehavi and Starr (2014) reports a racial disparity unfavorable to Black defendants in being charged with offenses carrying mandatory minimums, the federal equivalence of sentence enhancement in North Carolina. Specifically, federal prosecutors are 1.75 times more likely to charge Black defendants with offenses carrying mandatory minimum than the white counterparts. It is worth highlighting that Rehavi and Starr (2014) examines federal prosecutors’ mandatory minimum charging decisions, while we analyze defendants’ take-up sentence enhancements, which are judges’ decisions and conditional on state prosecutors’ initiation and valid proof of relevant aggravating factors.

Under the theoretical prediction in Proposition 3, these results provide no evidence that there is a racial bias against Black or Hispanic defendants underlying prosecutors’ decisions to pursue sentence enhancements. The event study plot shown in Figure 2.7 further corroborates the null finding of the effects of Blakely on racial disparity in receiving sentence enhancements. The figure also suggests an absence of pre-existing trends in racial disparities. All of the period-specific coefficients have small magnitude and they are not statistically significant at any conventional level.

VIII Mechanism: Prosecutors’ Choice

Our main analysis focuses on how Blakely affects defendants’ take-up of sentence enhancements in North Carolina. Sentence enhancement take-up is formally determined by judges, but contingent upon prosecutors’ requests of the enhancements and valid proof of relevant aggravating factors. Our analysis demonstrates, through a regression discontinuity approach, that Blakely substantially decreased defendants’ take-up of sentence enhancements. There are three plausible channels that may be driving our results. The first channel is that prosecutors responded to a higher burden of proof by deciding not to pursue sentence enhancements in a set of marginal cases where the benefits of pursuing the enhancements outweigh the costs, which is explained in our theoretical model in Section IV. We argue that this is the main channel that explains the estimated effects of Blakely on defendants’ sentence enhancements.

The second channel is the constraint on judicial discretion. Prior to the intervention, North Carolina judges were not only responsible for finding facts of aggravating factors but also determined whether to enhance sentences. Blakely took away judges’ authority of finding aggravating factors
and gave the fact-finding authority to juries, which constrained judicial discretion in determining sentence enhancements. Therefore, judges only make the final determination on whether to aggravate defendants’ sentences after *Blakely*. The new constraint on judicial discretion is unlikely to drive the main results for three reasons. First, *Blakely* did not change the prior rule that judges determine sentence enhancements by weighing aggravating factors against mitigating factors. Second, the North Carolina state law does not grant judges the departure power (*Wright* [2002]). The law does not allow judges to sentence defendants from aggregated ranges unless prosecutors request sentence enhancements and meet of burden of proving the relevant aggravating factors. Under this sentencing practice, defendants’ take-up of sentence enhancements are primarily attributable to prosecutorial discretionary decisions of seeking to enhance defendants’ sentences. Third, even if judges do respond to this new constraint, they would be more likely to impose sentence enhancements by weighing more heavily on any present aggravating factors, which is in the opposite direction of our main effects.

Alternatively, the effects of *Blakely* could also come from the discretion of jurors’ fact-finding. It is possible that, compared with judges, juries are less likely to find aggravating factors because they apply the beyond a reasonable doubt standard, rather than the preponderance of evidence standard (this is henceforth referred to as “jury effects”). There is both anecdotal and empirical evidence suggesting that the reduction in sentence enhancements are unlikely to be driven by jury effects.

Anecdotal evidence suggests that jury effects are minimal. First, when amending the state laws, North Carolina legislators projected that very few cases would trigger the new jury requirement, which could be easily accommodated by the court system (*Wright* [2005]). In addition, criminal defence attorneys responded to *Blakely* by filing motions to request jury trials if state prosecutors had sought aggravated sentences against the represented defendants. As a result, state prosecutors reversed their decisions to charge defendants with sentence enhancements (*Wright* [2005]).

Then, we examine the effects of *Blakely* on sentence enhancements using a subsample of criminal cases resolved through plea bargains. Table 2.1 shows that 86% of cases in our sample are resolved through plea bargains. The share of plea bargain cases increases to 96% if we exclude the fraction of cases that are dismissed voluntarily by prosecutors before moving to the plea bargain stage. Among cases resolved by plea bargains, defendants can only receive sentence enhancements if prosecutors intend to seek for the enhancements and defendants admit to the relevant aggravating factors. Juries do not play any role in determining defendants’ take-up of sentence enhancements, which
mechanically isolate the “jury effects.”

One might be concerned that the Blakely may have endogenously changed the composition of criminal cases resolved through plea bargains. Such concern is plausible as Blakely is likely to reduce prosecutors’ expected payoffs for taking cases to trials because the more stringent fact-finding standard made it more difficult to punish defendants with sentence enhancements. As a result, one might expect that Blakely leads to an increase in the share of criminal cases resolved by plea bargains. However, we do not find any empirical evidence suggesting that Blakely leads to any significant change in likelihood of cases being resolved by plea bargains. Figure 2.B.2 presents graphic evidence that Blakely did not lead to a discontinuous change in the share of cases resolved by plea bargains. Table 2.C.6 shows the regression results. The outcome variable is a dummy indicating whether a case is resolved through plea bargain. All of the point estimates have small magnitude and are not statistically significant, which corroborates the graphical evidence in Figure 2.B.2. Overall, the evidence does not suggest that Blakely significantly changed the composition of cases resolved through plea bargains.

Table 2.C.7 shows the results for the effects of Blakely on sentence enhancements among cases resolved through plea bargains. All of the point estimates are statistically significant at 1% level. Based on the Column (3) specification, Blakely leads to a 1.9 percentage point decrease, 61% decrease from the pre-treatment mean, in the likelihood of receiving sentence enhancements among defendants whose cases are resolved by plea bargains. The estimates are almost identical to our findings from whole sample. Mechanically, the results are not driven by juries applying a more stringent fact-finding standard. Instead, it is the prosecutors’ behavioral changes that primarily explain the decrease in defendants’ take-up of sentence enhancements.

IX Results for the Effects of Blakely on Gender Disparity in Sentence Enhancements by Charging Offense Type

Section VII.2 shows that Blakely closed the preexisting gender gap in sentence enhancements. The gap closure is driven by two main plausible channels: gender differences in the unobserved gender differences in aggravating factors that affect the take-up of sentence enhancements and judges and prosecutors’ preferences to disproportionately target men for sentence enhancement. In this section, we examine the effects of Blakely on gender disparity in sentence enhancements using sub-samples of defendants with the same charging offenses. To a large extent, such empirical
approach mitigates the unobserved gender differences in aggravating factors because defendants with the same charging offense are likely to receive enhancements on similar grounds. Therefore, the empirical analysis sheds more lights on prosecutors’ gender bias against men in determining whether to pursue sentence enhancements against defendants.

Figure 2.8 plots the offense-specific coefficients, with 95% confidence intervals, for the interaction term, 1[Blakely] * Male, in the main empirical model. The estimates are ordered based on the frequencies of charging offenses. Among defendants charged with the same felony offense, Blakely disproportionately benefited men by causing a larger decrease in the likelihood of receiving sentence enhancements except those charged with forgery, homicide, and embezzlement. Among defendants charged with forgery, Blakely made men significantly more likely to receive sentence enhancements than women.

We further examine the effects of Blakely on gender disparity in sentence enhancements among defendants charged with assault and forgery. Figure 2.8 shows the effects of Blakely on gender disparity in sentence enhancements are opposite among these two groups of defendants. In North Carolina, felony assault is a violent crime and typically involves at least one of the following elements: use of deadly weapons, inflicting serious injuries, or an intent to kill. In contrast, forgery is non-violent and defendants typically face forgery charges if they counterfeit instruments, security, or endorsement with a clear intent to defraud individuals, financial institutions or governmental units. Table 2.2 shows that assault and forgery account for 6% and 3% of felony cases in our sample. Furthermore, there is a larger share of women among defendants charged with forgery than those facing assault charges. Table 2.C.2 shows that 44.3% of defendants charged with forgery are women, almost 4 times as large as the fraction of female defendants among those charged with assault, 11.7%. Also, defendants charged with assault are more than twice as likely to receive sentence enhancements as those charged with forgery.

Table 2.6 shows the results for the effects of Blakely on gender disparity in sentence enhancements among defendants charged with felony assault. Once again, estimates from our preferred specification are given in Column (5). The estimated coefficient for the male dummy is 0.023 and it is statistically significant at the 1% level. This implies that pre-Blakely, men charged with assault

79 More than 80% of felony assault cases are brought under the following three statutory provisions: (1) § 14-32 Felonious Assault with Deadly Weapon with Intent to Kill or Inflicting Serious Injury; (2) § 14-32.4. Assault inflicting serious bodily injury; strangulation; penalties, and (3) § 14-34.2. Assault with a firearm or other deadly weapon upon governmental officers or employees, company police officers, or campus police officers

80 More than 95% of the forgery cases in the sample are brought under the following two statutory provisions: (1) § 14-119. Forgery of notes, checks, and other securities; counterfeiting of instruments; (2) § 14-120. Uttering forged paper or instrument containing a forged endorsement.
are 2.3 percentage points, or 71%, more likely to receive sentence enhancements than women. The estimated coefficient for $1[\text{Blakely}]$ is -0.001 and not statistically significant, which implies that \textit{Blakely} did not lead to meaningful changes in women’s likelihood of receiving sentence enhancements. In contrast, the point estimate of the interaction term, $1[\text{Blakely}] \times \text{Male}$, is -0.023 and it is statistically significant at 1% level. This result suggests that \textit{Blakely} leads to a 2.3 percentage point larger decrease in male defendants’ likelihood of receiving sentence enhancements, which closes almost the entire pre-existing gender gap.

We repeat this analysis for the subset of cases with forgery charges. The results are reported in Table 2.7. Surprisingly, forgery cases have effects moving in the opposite direction from our main findings. First, we find a pre-\textit{Blakely} gender gap in sentence enhancements unfavorable to women. In Column (5), the point estimate of the male dummy is -0.009 and statistically significant at 10% level, which suggests a pre-\textit{Blakely} gender gap with women being 0.9 percentage points, or 43%, more likely to receive sentence enhancements than men. Second, \textit{Blakely} disproportionately benefited women charged with forgery in terms of avoiding sentence enhancements. The point estimates for $1[\text{Blakely}]$ and $1[\text{Blakely}] \times \text{Male}$ are -0.02 and 0.017 respectively, statistically significant at 1% level. The estimates imply that, among defendants charged with forgery, \textit{Blakely} made men and women 0.3 and 2 percentage points less likely to receive sentence enhancements. These results are quite striking. It means that \textit{Blakely} not only closed the pre-existing gender gap unfavorable to women but also made men substantially more likely to receive sentence enhancements than women.

It is highly likely that the opposite gender disparity results among the two subgroups of defendants are driven by the unobserved aggravating factors that affect the take-up of sentence enhancements. Assault is a violent crime and defendants’ sentence enhancements are likely to result from hate crimes or being “heinous, atrocious or cruel.” It is likely that \textit{Blakely} disproportionately benefited male defendants charged with assault because prosecutors tend to associate male defendants with hate crimes or cruelty more than they do for women. These factors are difficult to establish beyond a reasonable doubt compared to other factors that are more straightforward to prove. In contrast, forgery is not a violent crime, and defendants charged with forgery are unlikely to receive sentence enhancements due to reasons related to violence. Instead, taking advantage of “a position of trust or confidence” is the most likely aggravating factor that enhances the sentences of defendants convicted of forgery. The most plausible reason why \textit{Blakely} disproportionately benefited women charged with forgery is that prosecutors associate women with taking advantage of the relationship of trust and confidence and therefore pursue sentence enhancements when the evidence
is weak. Overall, the contrasting results among defendants charged with felony assault and forgery imply that prosecutors tend to associate defendants’ gender with different types of aggravating factors in deciding whether to pursue sentence enhancements.

**X Conclusion**

Prosecutors in the US exercise a wide range of discretion in determining how to process criminal cases. Understanding the implication of this exercise of discretion offers valuable insights for future institutional reform of constraining or expanding prosecutorial discretion. Our paper evaluates the causal effects of restricting prosecutorial discretion on racial and gender disparity, which often raises public concerns over constitutional equal protection. The empirical analysis evaluates a natural experiment introduced by *Blakely v. Washington* which restricted North Carolina state prosecutors’ discretion by requiring them to meet a higher burden of proof standard in order to pursue sentence enhancements.

In this paper, we have provided a conceptual model of how burdens of proof affect prosecutors’ decisions to charge defendants with sentence enhancements. The model predicts that a higher burden of proof standard makes prosecutors less likely to pursue sentence enhancements due to higher costs and lower expected benefits. The model also predicts that the decrease in sentence enhancements will be larger in subgroups of defendants being discriminated by prosecutors. We employ a regression discontinuity framework to exploit the timing of the *Blakely* decision as a natural experiment in shifting to a higher burden of proof standard. We find that imposing a higher burden of proof decreases defendants’ likelihood of receiving sentence enhancements by 66%, which is largely driven by prosecutors’ behavioral changes.

Examining the discontinuity across different racial and gender groups, we find a large unexplained gender gap with men being more likely to receive sentence enhancements than women after controlling extensive sets controls and fixed effects. The entire unexplained gender gap of sentence enhancements was closed after requiring prosecutors to prove relevant aggravating factors by beyond a reasonable doubt. In contrast, we find no evidence of racial disparity unfavorable to Black or Hispanic defendants. There is also no evidence that *Blakely* leads to any differential impacts on Black/Hispanic defendants.

We further examine how *Blakely* affects gender disparity in sentence enhancements among defendants with the same charging offense. Defendants with the same charging offenses are likely to
receive sentence enhancements for similar reasons, which mitigates the concerns that the main results of gender disparity are driven by gender differences in the unobserved aggravating factors. We find that Blakely disproportionately benefited men charged with assault in reducing the likelihood of receiving sentence enhancements. However, among defendants charged with forgery, it is women who disproportionately benefited from requiring prosecutors to prove aggravating factors beyond a reasonable doubt. This stark contrast suggests that prosecutors tend to associate defendants’ gender with different types of aggravating factors and therefore pursue sentence enhancements when the evidence is not sufficient to survive proof by beyond a reasonable doubt.

The Supreme Court’s development of Sixth Amendment jurisprudence caused important institutional changes in the criminal justice system. Empirically evaluating these changes with data and econometric methods often yields valuable evidence which informs socially optimal policy decisions. This paper makes an important step in evaluating how the Court’s enforcement of the defendants’ right to a jury trial affects state prosecutors’ behavior in North Carolina. Federal prosecutors might behave differently from their state counterparts as federal prosecutors are selected by appointment, instead of democratic elections (Stuntz 2001). Therefore, our findings on racial and gender disparities might not speak to how federal prosecutors behave. More future research is encouraged in this direction.
** Effective for Offenses Committed on or after 12/1/95 Through 11/30/09 **

### FELONY PUNISHMENT CHART

<table>
<thead>
<tr>
<th>PRIOR RECORD LEVEL</th>
<th>I 0 Points</th>
<th>II 1-4 Points</th>
<th>III 5-8 Points</th>
<th>IV 9-14 Points</th>
<th>V 15-18 Points</th>
<th>VI 19+ Points</th>
<th>DISPOSITION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DISPOSITION</strong></td>
<td><strong>DEATH OR LIFE WITHOUT PAROLE</strong></td>
<td><strong>LIFE WITHOUT PAROLE</strong></td>
<td><strong>LIFE WITHOUT PAROLE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>240 - 300</td>
<td>288 - 360</td>
<td>336 - 420</td>
<td>384 - 480</td>
<td>346 - 433</td>
<td>384 - 480</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>73 - 75</td>
<td>100 - 125</td>
<td>116 - 145</td>
<td>133 - 167</td>
<td>151 - 188</td>
<td>151 - 188</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>64 - 66</td>
<td>77 - 95</td>
<td>103 - 129</td>
<td>117 - 146</td>
<td>133 - 167</td>
<td>133 - 167</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>16 - 20</td>
<td>19 - 24</td>
<td>21 - 26</td>
<td>26 - 36</td>
<td>32 - 42</td>
<td>32 - 42</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>13 - 16</td>
<td>15 - 19</td>
<td>17 - 21</td>
<td>20 - 25</td>
<td>24 - 31</td>
<td>24 - 31</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>6 - 8</td>
<td>8 - 10</td>
<td>10 - 12</td>
<td>11 - 14</td>
<td>15 - 19</td>
<td>15 - 19</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>4 - 6</td>
<td>6 - 8</td>
<td>8 - 10</td>
<td>9 - 12</td>
<td>12 - 16</td>
<td>12 - 16</td>
<td></td>
</tr>
</tbody>
</table>

**OFFENSE CLASS**

<table>
<thead>
<tr>
<th>A – Active Punishment</th>
<th>I – Intermediate Punishment</th>
<th>C – Community Punishment</th>
</tr>
</thead>
</table>

Note: The figure shows the change in defendants’ likelihood of receiving sentence enhancements at the *Blakely* discontinuity (June 2004). The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of the cases charged with felony offenses in North Carolina from 2001 to 2007.
Figure 2.3: Monthly Means of Defendants Receiving sentence enhancements (By Gender)

Note: The figure shows the gender-specific changes in the likelihood of receiving sentence enhancements at the Blakely discontinuity (June 2004). The curves are fitted on each side of Blakely with a second-degree polynomial function of the time. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of the cases charged with felony offenses in North Carolina from 2001 to 2007.
Figure 2.4: Monthly Share of Defendants Receiving sentence enhancements (By Race)

Note: The figure shows the race-specific change in the likelihood of receiving sentence enhancements at the Blakely discontinuity (June 2004). The curves are fitted on each side of Blakely with a second-degree polynomial function of the time. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of the cases charged with felony offenses in North Carolina from 2001 to 2007.
Figure 2.5: Changes of Case Volumes at the *Blakely* Discontinuity

Note: The figure shows the McCrary test results on the number of cases at the *Blakely* discontinuity. The curves are fitted on each side of *Blakely* with a second-degree polynomial function of the time. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of the cases charged with felony offenses in North Carolina from 2001 to 2007.
Figure 2.6: Event Study Plot: Gender Disparity in Sentence Enhancements

Note: The figure shows the event-study estimates of gender disparity in sentence enhancements, with the 95% confidence intervals. The regression model is as follows:

\[ Aggravated_{ijkc} = \beta_0 + \beta_1 \text{Male}_i + \beta_2 \text{Blakely}_i + \sum_t \eta_t (\text{Male}_i \times \text{Blakely}_t) + X_{ij} \gamma + Grid_{jk} + \theta_c + \epsilon_{ijkc} \]

Time is measured at the semi year level. The regression specification include defendants’ demographics, guidelines fixed effects and county fixed effects. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level.
Figure 2.7: Event Study Plot: Racial Disparity in Sentence Enhancements

Note: The figure shows the event-study estimates of gender disparity in sentence enhancements, with the 95% confidence intervals. The regression model is as follows:

\[ \text{Aggravated}_{ijkc} = \beta_0 + \beta_1 \text{Black/Hispanic}_i + \beta_2 [\text{Blakely}] + \sum_t \eta_t (\text{Black/Hispanic}_i \times [\text{Blakely}_t]) + X_i \gamma + \text{Grid}_{jk} + \theta_c + \epsilon_{ijkc} \]

Time is measured at the semi year level. The regression specification include defendants' demographics, guidelines fixed effects and county fixed effects. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level.
Figure 2.8: Offense-Specific Changes in Gender Disparity of Sentence Enhancements After *Blakely*

Note: We separately run the following regression specification on each type of charged felony offenses:

\[ Aggravated_{ijkcm} = \beta_0 + \beta_1 \text{Male}_i + \beta_2 1[\text{Blakely}] + \beta_3 1[\text{Blakely}] \ast \text{Male} + X_i \gamma + Grid_{jk} + \theta_c + \delta_m + \epsilon_{ijkcm} \]

The figure plots offense-specific estimates of $\beta_3$ in local linear regression. The estimates are ordered from top to bottom based the number of observations for each offense. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level.
Table 2.1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.84</td>
<td>0.37</td>
</tr>
<tr>
<td>White</td>
<td>0.40</td>
<td>0.49</td>
</tr>
<tr>
<td>Black/Hispanic</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td>Age</td>
<td>30.10</td>
<td>10.35</td>
</tr>
<tr>
<td>Private Attorney</td>
<td>0.24</td>
<td>0.43</td>
</tr>
<tr>
<td>Public Defender</td>
<td>0.22</td>
<td>0.41</td>
</tr>
<tr>
<td>Criminal History Points</td>
<td>4.60</td>
<td>5.30</td>
</tr>
<tr>
<td>Arrest Rate (County)</td>
<td>535.96</td>
<td>221.07</td>
</tr>
<tr>
<td>Unemployment Rate (County)</td>
<td>5.81</td>
<td>1.57</td>
</tr>
<tr>
<td>Plea Bargain</td>
<td>0.86</td>
<td>0.34</td>
</tr>
<tr>
<td>Plead Guilty to Lower Offenses</td>
<td>0.24</td>
<td>0.43</td>
</tr>
<tr>
<td>Sentence Enhancement</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>Mitigated</td>
<td>0.08</td>
<td>0.27</td>
</tr>
<tr>
<td>Minimum Sentence Length</td>
<td>17.37</td>
<td>39.45</td>
</tr>
<tr>
<td>Incarceration</td>
<td>0.31</td>
<td>0.46</td>
</tr>
<tr>
<td>Observations</td>
<td>213264</td>
<td></td>
</tr>
</tbody>
</table>

Note: The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). Arrest rate is defined as the number of arrests made per 100,000 county population. The sample includes cases charged with felony offenses from 2001 to 2007. 95.9% of the cases are resolved by plea bargain if we drop the cases voluntarily dismissed by prosecutors.
Table 2.2: Charged Felony Offenses and Shares

<table>
<thead>
<tr>
<th>Offense</th>
<th>Frequency</th>
<th>Share (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assault</td>
<td>12553</td>
<td>5.89</td>
</tr>
<tr>
<td>Burglary</td>
<td>25432</td>
<td>11.93</td>
</tr>
<tr>
<td>Drugs</td>
<td>66994</td>
<td>31.41</td>
</tr>
<tr>
<td>Embezzlement</td>
<td>3203</td>
<td>1.50</td>
</tr>
<tr>
<td>Forgery</td>
<td>6609</td>
<td>3.10</td>
</tr>
<tr>
<td>Fraud</td>
<td>16745</td>
<td>7.85</td>
</tr>
<tr>
<td>Homicide</td>
<td>3359</td>
<td>1.58</td>
</tr>
<tr>
<td>Kidnapping</td>
<td>2345</td>
<td>1.10</td>
</tr>
<tr>
<td>Larceny</td>
<td>24524</td>
<td>11.50</td>
</tr>
<tr>
<td>Robbery</td>
<td>13670</td>
<td>6.41</td>
</tr>
<tr>
<td>Sexual Assault</td>
<td>10907</td>
<td>5.11</td>
</tr>
<tr>
<td>Traffic</td>
<td>9597</td>
<td>4.50</td>
</tr>
<tr>
<td>Weapon Offense</td>
<td>6898</td>
<td>3.23</td>
</tr>
<tr>
<td>Other</td>
<td>10428</td>
<td>4.89</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>213264</strong></td>
<td><strong>100.00</strong></td>
</tr>
</tbody>
</table>

Note: The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS).
Table 2.3: The Effects of Blakely on Sentence Enhancements (Overall)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1[\text{Blakely}]$</td>
<td>-0.0157***</td>
<td>-0.0180***</td>
<td>-0.0187***</td>
<td>-0.0157***</td>
<td>-0.0185***</td>
</tr>
<tr>
<td></td>
<td>(0.00112)</td>
<td>(0.00126)</td>
<td>(0.00127)</td>
<td>(0.00189)</td>
<td>(0.00127)</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Month-of-Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Guidelines Fixed Effects</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local Linear Regression</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Polynomial Regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>213264</td>
<td>186825</td>
<td>186825</td>
<td>186825</td>
<td>186825</td>
</tr>
</tbody>
</table>

Note: The pre-Blakely mean of sentence enhancements is 0.029. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of criminal cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level. Significance Levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table 2.4: The Effects of *Blakely* on Gender Disparity in Sentence Enhancements

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>-0.00927***</td>
<td>-0.0107***</td>
<td>-0.0115***</td>
<td>-0.00885***</td>
<td>-0.0117***</td>
</tr>
<tr>
<td></td>
<td>(0.00175)</td>
<td>(0.00199)</td>
<td>(0.00199)</td>
<td>(0.00254)</td>
<td>(0.00196)</td>
</tr>
<tr>
<td>Male</td>
<td>0.0124***</td>
<td>0.00780***</td>
<td>0.00777***</td>
<td>0.00778***</td>
<td>0.00705***</td>
</tr>
<tr>
<td></td>
<td>(0.00133)</td>
<td>(0.00149)</td>
<td>(0.00148)</td>
<td>(0.00148)</td>
<td>(0.00147)</td>
</tr>
<tr>
<td>1[Blakely] * Male</td>
<td>-0.00770***</td>
<td>-0.00872***</td>
<td>-0.00849***</td>
<td>-0.00848***</td>
<td>-0.00807***</td>
</tr>
<tr>
<td></td>
<td>(0.00160)</td>
<td>(0.00180)</td>
<td>(0.00178)</td>
<td>(0.00178)</td>
<td>(0.00175)</td>
</tr>
</tbody>
</table>

Controls          | X  | X  | X  | X  |
County Fixed Effects |   |   |   |   |
Month-of-Year Fixed Effects | X | X | X | X |
Guidelines Fixed Effects |   |   |   |   |
Local Linear Regression | X | X | X | X |
Polynomial Regression | X  |   |   |   |
N                | 213264 | 186825 | 186825 | 186825 | 186825 |

Note: The pre-Blakely mean of sentence enhancements is 0.029. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of criminal cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 2.5: The Effects of *Blakely* on Racial Disparities in Sentence Enhancements

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>-0.0155***</td>
<td>-0.0177***</td>
<td>-0.0184***</td>
<td>-0.0157***</td>
<td>-0.0181***</td>
</tr>
<tr>
<td></td>
<td>(0.00138)</td>
<td>(0.00158)</td>
<td>(0.00158)</td>
<td>(0.00212)</td>
<td>(0.00159)</td>
</tr>
<tr>
<td>Black/Hispanic</td>
<td>0.000921</td>
<td>-0.00143</td>
<td>-0.00130</td>
<td>-0.00130</td>
<td>-0.00147</td>
</tr>
<tr>
<td></td>
<td>(0.00113)</td>
<td>(0.00125)</td>
<td>(0.00123)</td>
<td>(0.00123)</td>
<td>(0.00124)</td>
</tr>
<tr>
<td>1[Blakely] * Black/Hispanic</td>
<td>-0.000373</td>
<td>-0.000458</td>
<td>-0.000493</td>
<td>-0.000495</td>
<td>-0.000763</td>
</tr>
<tr>
<td></td>
<td>(0.00128)</td>
<td>(0.00141)</td>
<td>(0.00140)</td>
<td>(0.00140)</td>
<td>(0.00140)</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Month-of-Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guidelines Fixed Effects</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local Linear Regression</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polynomial Regression</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>213264</td>
<td>186825</td>
<td>186825</td>
<td>186825</td>
<td>186825</td>
</tr>
</tbody>
</table>

Note: The pre-*Blakely* mean of sentence enhancements is 0.029. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of criminal cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 2.6: The Effects of *Blakely* on Gender Disparity in Sentence Enhancements (Assault)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>-0.00428</td>
<td>-0.00251</td>
<td>-0.00454</td>
<td>-0.00941</td>
<td>-0.00119</td>
</tr>
<tr>
<td></td>
<td>(0.00765)</td>
<td>(0.00785)</td>
<td>(0.00787)</td>
<td>(0.00820)</td>
<td>(0.0107)</td>
</tr>
<tr>
<td>Male</td>
<td>0.0286***</td>
<td>0.0292***</td>
<td>0.0224***</td>
<td>0.0231***</td>
<td>0.0232***</td>
</tr>
<tr>
<td></td>
<td>(0.00624)</td>
<td>(0.00597)</td>
<td>(0.00633)</td>
<td>(0.00652)</td>
<td>(0.00651)</td>
</tr>
<tr>
<td>1[Blakely] * Male</td>
<td>-0.0195***</td>
<td>-0.0226***</td>
<td>-0.0209***</td>
<td>-0.0225***</td>
<td>-0.0225***</td>
</tr>
<tr>
<td></td>
<td>(0.00696)</td>
<td>(0.00686)</td>
<td>(0.00710)</td>
<td>(0.00747)</td>
<td>(0.00747)</td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Month-of-Year Fixed Effects</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Guidelines Fixed Effects</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Local Linear Regression</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Polynomial Regression</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>12553</td>
<td>11334</td>
<td>11334</td>
<td>11334</td>
<td>11334</td>
</tr>
</tbody>
</table>

Note: The pre-*Blakely* mean of sentence enhancements is 0.044 among defendants charged with felony assault. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of criminal cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01.
Table 2.7: The Effects of Blakely on Gender Disparity in Sentence Enhancements (Forgery)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>-0.0181***</td>
<td>-0.0252***</td>
<td>-0.0261***</td>
<td>-0.0276***</td>
<td>-0.0223**</td>
</tr>
<tr>
<td></td>
<td>(0.00558)</td>
<td>(0.00757)</td>
<td>(0.00761)</td>
<td>(0.00885)</td>
<td>(0.00969)</td>
</tr>
<tr>
<td>Male</td>
<td>-0.00332</td>
<td>-0.00855*</td>
<td>-0.00908*</td>
<td>-0.00904*</td>
<td>-0.00906*</td>
</tr>
<tr>
<td></td>
<td>(0.00385)</td>
<td>(0.00495)</td>
<td>(0.00488)</td>
<td>(0.00488)</td>
<td>(0.00488)</td>
</tr>
<tr>
<td>1[Blakely] * Male</td>
<td>0.0127**</td>
<td>0.0165***</td>
<td>0.0159**</td>
<td>0.0169***</td>
<td>0.0168***</td>
</tr>
<tr>
<td></td>
<td>(0.00490)</td>
<td>(0.00616)</td>
<td>(0.00606)</td>
<td>(0.00603)</td>
<td>(0.00603)</td>
</tr>
</tbody>
</table>

Controls X X X X X
County Fixed Effects X X X
Month-of-Year Fixed Effects X X X
Guidelines Fixed Effects X
Local Linear Regression X X X
Polynomial Regression X
N 6609 5209 5209 5209 5209

Note: The pre-Blakely mean of sentence enhancements is 0.0164 among defendants charged with forgery. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of criminal cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Chapter 3

The Direct and Spillover Effects of Enforcing the Element Rule
I. Introduction

In the early 2000s, the Supreme Court decided a series of criminal law cases, clarifying one crucial question: did the Sixth Amendment right to a jury trial apply to fact findings that enhance the length of defendants’ sentence? The Court’s decisions elaborated on an “element rule,” meaning that the sentence enhancement facts are “elements of the crime” and only juries can find those facts by beyond a reasonable doubt. Before the Supreme Court cases, sentence enhancement facts are sentencing factors which judges decide by a lower standard of preponderance of evidence. Courts’ enforcement of the element rule resulted in a massive reform, if not a complete overhaul, of the mandatory sentencing guideline systems at both the federal and state levels.

The element rule is intended to safeguard defendants’ legitimate interests against unwarranted punishment. Apprendi v. New Jersey is the earliest Supreme Court case that enforced the element rule. The Court held that facts increasing statutory maximum sentences should be proved to juries “beyond a reasonable doubt.” The controversial decision triggered an academic debate on whether the element rule can realistically protect defendants’ interests in the criminal justice system, in which discretion is often exercised at many points in the case.

Stefanos Bibas provocatively argued that the element rule is harmful to defendants because it empowers federal prosecutors, particularly “in the world guilty plea.” The element rule is likely to trigger more aggressive prosecution, making defendants forgo important legal rights and become more susceptible to sentence enhancement. Nancy King and Susan Klein dismissed Bibas’s argument as “indeed startling” but “dead wrong.” They argued that the element rule not only makes it more difficult for prosecutors to prove sentence enhancement facts but also deters prosecutors from forcing defendants to admit to those facts in plea bargains. This debate has been influential and affected the Court’s decision in the most consequential right-to-jury case, Blakely v. Washington. In Blakely, the Supreme Court Justices debated on the very issue of whether the element rule harms the interests of criminal defendants. Justice Scalia’s majority opinion cited King & Klein and Justice Breyer’s dissent opinion citing Bibas as persuasive authorities.

Ultimately, whether the element rule is harmful to criminal defendants is an empirical question. The answers must rely on data and rigorous empirical tests. Also, the answers might vary with jurisdictions with different procedural rules. To my knowledge, no scholarly efforts have performed such empirical work. This article uses administrative data from North Carolina to fill
the gap in the literature. In addition, I propose a theoretical framework on how prosecutors decide to pursue sentence enhancement, which incorporates the seemingly conflicting views between Bibas and King & Klein.

Under the model, prosecutors are assumed to maximize the net benefit of punishing criminal defendants. The model predicts a “direct effect” in which enforcing the element rule makes defendants less likely to receive sentence enhancements because enforcing the rule increases the costs and decreases the expected benefits of pursuing sentence enhancements. The “direct effect” matches the reasoning by King & Klein. On the other hand, enforcing the element rule may also result in unintended “spillover effects” due to prosecutors’ broad discretion through the entire criminal process. As Bibas suggested, prosecutors may plausibly compensate for the loss of opportunities to punish defendants with sentence enhancements by being more stringent against defendants in making other discretionary decisions.

Then, I empirically estimated the direct and spillover effects of enforcing the element rule using a large administrative dataset that consists of more than 200,000 felony cases disposed of in North Carolina superior courts. North Carolina sentencing follows a mandatory guideline system that prescribes presumptive, mitigated, and aggravated ranges for minimum sentences. I can measure sentence enhancement by examining whether the defendants’ minimum sentence comes from the aggravated range. The Court’s Blakely decision raised the evidentiary standard of proving sentence enhancement facts from “preponderance of evidence” to “beyond a reasonable doubt.” I used the Blakely decision as a natural experiment and employed a regression discontinuity framework to estimate the causal effects of enforcing the element rule on defendants’ case outcomes.

Narratives of the changes occurring after the Blakely show that the decision led to substantial change court activities immediately. I find strong evidence that enforcing element rule significantly reduced defendants’ likelihood of receiving sentence enhancements by roughly two percentage points, a 66% decrease from the pre-Blakely mean, which is a large effect. I found a similar result after restricting the sample to cases resolved by plea bargains. Therefore, it appeared that enforcing the element rule favored defendants by reducing the likelihood that their sentences would be enhanced, similar to King & Klein’s prediction.

Further, I estimated the spillover effects of Blakely by examining whether enforcing the element rule triggers prosecutors’ behavioral changes in making other prosecution decisions. The
results validated Bibas’s theory that prosecutors are likely to treat defendants more strictly in response to the new procedural barrier of punishing defendants with sentence enhancements. Indeed, *Blakely* increased the number of charges by 4.8%, the likelihood of defendants pleading to a lower felony by 8.3%, and the probability of receiving incarceration punishment in a plea bargain by 4.3%. In addition, my analysis provides new insights that prosecutors also responded to *Blakely* by concentrating resources towards prosecuting more serious crimes. *Blakely* increased the cases dismissal rate by 15.8% and the likelihood of defendants pleading from a felony to a misdemeanor by 5.4%. Finally, I did not find evidence suggesting that *Blakely* leads to significant changes in defendants’ minimum sentence lengths. The critical insight from the spillover effects analysis is that prosecutors’ exercise of broad discretion can undermine the potential benefits of the element rule in protecting criminal defendants from unjust punishment.

Another contribution is the empirical finding that constraining prosecutorial discretion reduced unwarranted disparities. I examined how *Blakely* changed the unwarranted disparities at two levels, inter-district and within-district. The inter-district disparity is the unexplained gap in sentence enhancement rates between the most lenient and the most stringent district in a given month. The within-district disparity is the unexplained gap in sentence enhancements rate between the most stringent and least stringent months within a district. My analysis shows that *Blakely* significantly reduced both inter-district and within-district differentials. The finding revealed that more than 50% of the pre-*Blakely* inter-district and within-district disparities of sentence enhancements are attributable to the low preponderance of evidence standard in enhancing defendants’ sentence lengths.

The rest of the article is structured as follows: Part 2 reviews the academic debate on whether the element rule is harmful to criminal defendants. In this part, I also develop a new conceptual framework that incorporates both of the competing views. Part 3 reviews the post-*Apprendi* case laws on the Court’s enforcement of the element rule and how the new Sixth Amendment Jurisprudence changed the sentencing practice at federal and state levels. Part 4 reviews the current empirical research to date on prosecutors’ discretion. Part 5 reviews the relevant institutional features of North Carolina sentencing, the context of my empirical analysis. Part 6 describes the data, summary statistics, and the empirical frameworks. Part 7 describes the results for the direct and spillover effects of *Blakely*. Part 8 describes the results of how *Blakely*
affected the inter-district and within-district disparities of sentence enhancements. Part 9 concludes with some remarks.

II. Whether Right to Jury Protects Criminal Defendants

Until 2000, the Supreme Court’s right to jury jurisprudence did not offer a coherent answer to whether sentence enhancement facts need to be found by a jury beyond a reasonable doubt. Before 2000, the Court appeared to permit legislatures to define sentencing factors that judges may find by a preponderance of evidence standard. However, the Court shifted the course and began more rigorously enforcing jury guarantees from the *Apprendi v. New Jersey* decision. It held that any factor (other than prior convictions) that increases defendants’ sentencing length beyond the statutory maximum must be an element of crime that should be proved to a jury beyond a reasonable doubt, which is also defined as the “element rule.”

The *Apprendi* decision triggered a heated academic debate over the legitimacy of its ruling, which significantly affected the Court’s subsequent right to jury decisions. This paper focuses on one specific aspect of the debate, whether the element rule protects or harms defendants’ interests. This section begins by reviewing the *Apprendi* decision which provides an additional background of the debate. Then, I outline the legal academics’ competing theories on whether the element rule is beneficial or harmful to criminal defendants. Finally, I consolidate the conflicting theories from

---

1 On the one hand, in some cases, the Court vindicated defendants’ right to jury in finding facts which increase their criminal punishments. *In re Winship*, 397 U.S. 358 (1970) (The Court held that “every fact necessary to constitute the crime charged” needs to be found beyond reasonable doubt); *Mullaney v. Wilbur*, 421 U.S. 684 (1975) (The Court struck down a Maine state law which presumes defendants’ malice in defining murder crime and requires defendants to rebut the presumption by preponderance of evidence. It held that “malice” is an element of a crime, which juries should find by beyond reasonable doubt).

However, the Court’s other decisions appear to recognize judges’ broad discretion in making sentencing decisions and also the legislative power to define sentencing factors which do not trigger jury process. See *Williams v. New York*, 337 U.S. 241 (1949) (In determining if to sentence a defendant life imprisonment or death penalty, New York judges can legitimately exercise discretion to consider information from adverse witnesses who the defendant had not been permitted to confront or cross-examine). See *McMillan v. Pennsylvania*, 477 U.S. 79 (1986) (Upheld a Pennsylvania state criminal law which subject defendants to mandatory minimum sentence if judges find, by preponderance of evidence standard, that the person “visibly possess a firearm” in the course of committing a felony offense. For the first time, the Court invoked the distinction between “element of crime” and “sentencing factors.” The Court recognized legislative power to define sentencing factor and elaborated a five-factor balancing test to determine whether sentence enhancement facts are “elements of crime”); See *Almendarez-Torres v. United States*, 523 U.S. 224 (1998), at 224 (The Court upheld that a federal statute which allows sentence enhancement on the basis prior convictions not alleged in indictment).
Bibas and King & Klein by proposing a simple conceptual framework that predicts the direct and spillover effects of enforcing the element rule.

II.1. Apprendi v. New Jersey

In \textit{Apprendi}, defendant Charles Apprendi fired gunshots at an African American family who recently moved to his neighborhood that had been all white. Through a plea agreement, the defendant pleaded guilty over 3 of the 23 counts of initially indicted charges: count 3 and 18 of second-degree possession of firearms with unlawful purposes and count 23 of third-degree unlawful possession of an antipersonnel bomb.\footnote{530 U.S. 466 (2000)} The plea agreement stipulated that the prosecutors reserved the right to seek sentence enhancement for count 18 under the state “hate crime” statute and the defendant reserved the right to contest it.\footnote{N.J. Stat. Ann. § 2C:44-3(e) (The hate crime statute provision allows extended imprisonment if “the defendant in committing the crime acted with a purpose to intimidate an individual or group of individuals because of race, color, gender, handicap, religion, sexual orientation or ethnicity.”) [This section was amended in 2001]} Under state law, second-degree possession of firearms with unlawful purpose is punishable by imprisonment between 5 and 10 years. However, the state hate crime statute aggravates the sentencing range to 10 and 20 years.

The state prosecutors decided to move for a sentence enhancement for Count 18 offense due to their belief in the defendant’s racially biased motive. Then, the state court held a sentencing hearing where a state judge found the presence of the defendant’s racially biased motive by a preponderance of evidence standard. As a result, the defendant received a sentence of 12 years, exceeding the 10-year maximum sentence for the convicted offense.\footnote{Id. (The law explicitly authorized state judges to find relevant facts by preponderance of evidence standard.)} The defendant challenged the constitutionality of the hate crime enhancement procedure for not requiring prosecutors to prove his racially biased motives to a jury beyond a reasonable doubt.\footnote{Apprendi, at 475}

The main issue is whether the racially biased motive is an element of a crime or a sentencing factor. In a 5-4 decision, the Court ruled in favor of the defendant and held that any facts, other than prior convictions, that enhance the sentence beyond statutory maximum are “elements of crime,” which should be proved to a jury beyond a reasonable doubt.\footnote{Id., at 466} The Court dismissed the
state’s argument that the hate crime statute’s “purpose to intimate” is a simple inquiry of motive.\(^8\) Instead, the majority opinion argued that the law allowed a substantial sentence increase beyond the statutory maximum based on the defendant’s state of mind, or \textit{mens rea}, which triggers jury factfinding procedure under the Sixth Amendment.\(^9\)

\textit{Apprendi} has been generally viewed as a “watershed” case\(^10\) which marked a major change in the Court’s jury guarantee jurisprudence. As Justice O’Connor pointed out in the dissent opinion that, in the past 200 years or longer, the Court never constrained legislative power to define crimes with a “universal and seemingly bright-line rule.”\(^11\) \textit{Apprendi} leads to a series of subsequent cases which more vigorously enforced defendants’ right to jury and fundamentally transformed the sentencing system at both federal and state levels.

\textbf{II.2. Harmful to Defendant}

Stephanos Bibas provocatively argued that the element rule is detrimental to criminal defendants due to the broad prosecutorial power and the institutional loopholes. It is worth emphasizing that the argument is largely made in the context of prosecuting federal crime and the element rule is equivalent to the Court’s \textit{Apprendi} decision. The Court’s subsequent decisions further extended and enriched the rule far beyond the initial \textit{Apprendi} holding. Bibas questioned the legitimacy of the element rule by dismissing a variety of rationales underlying the Court’s \textit{Apprendi} decision, such as historic practice, preventing a slippery slope, and providing clear notice to defendants. But my empirical analysis directly speaks to his speculation that the element rule is detrimental to defendants in the “world of guilty pleas.”

Federal criminal defendants typically have two trials if they object to prosecutors’ plea offers – a jury or bench trial to determine conviction and a sentencing hearing to determine sentence length. Bibas argued that the element rule has quite limited benefits to defendants who choose to take their cases to trials. First, although the right to jury poses a stricter procedural barrier to find sentence enhancement facts, criminal defendants do not benefit from the procedural

\(^{8}\) \textit{Id}, at 492  
\(^{10}\) Barkow, Recharging Juries (cite)  
\(^{11}\) \textit{Apprendi}, at 525
protection unless they went up against sentences beyond the statutory maxima. What’s more, even if juries failed to find sentence enhancement facts, judges can still, at least theoretically, impose sentence enhancement by applying the preponderance of evidence standard to find the same facts. Therefore, such institutional loopholes compromised the benefits of the element rule.

The benefits of trial right further diminished in a “world of guilty pleas” where more than 95% of the federal criminal defendants are convicted through plea bargains. Federal prosecutors exercise broad power and dominate the process of charging and sentencing. Element rule incentivizes federal prosecutors to drive hard plea bargains, which leads to worse outcomes for criminal defendants. First, Bibas claimed the element rule tends to deprive defendants of the opportunity to have sentence hearings. In the pre-\textit{Apprendi} world, the defendant can choose to plead guilty only to the charged offenses and argue sentence enhancement facts in sentence hearings. However, the element rule pushes prosecutors to bargain more aggressively and force defendants to surrender to the sentence enhancement facts in plea agreements. Defendants therefore might lose the opportunity for sentencing hearings.

Also, prosecutors have broad discretionary authority to determine what charges to convict defendants and also whether to lower or even drop certain counts of charges, which is called “charge bargain.” The element rule essentially fragmented one single offense into multiple offenses based on statutory maxima, which undermines the judicial oversight of prosecutors. Bibas elaborates his argument using carjacking as an example. Under the federal carjacking statute, 18 U.S.C. § 2119, the maximum sentence for ordinary carjacking is 15 years, 25 years if the act results in “serious bodily injury” and life imprisonment if the act results in death. Under the old rule, judges are only constrained by the “life imprisonment” maximum and able to check whether prosecutors’ sentence decisions are consistent with facts. However, the element rule fragments carjacking into three separate offenses based on two sentence enhancement facts. If prosecutors decide to charge ordinary carjacking, judges are unable to raise sentences above 15 years even if “serious bodily injury” was clearly present.

Bibas’s argument had practical impacts on subsequent litigations. In June 2005, the Supreme Court decided \textit{Blakely v. Washington} which struck down the Washington state mandatory

\footnote{Barkow, Prisoners of Politics, Harvard University Press, P145 (Federal prosecutors in the department of justice not only set policies on charging and sentencing, but also make policy decisions related to forensics, corrections and clemency.)}
sentencing guideline for violating the Sixth Amendment.\textsuperscript{13} In the dissent opinion of \textit{Blakely}, Justice Breyer expressed deep fairness concern on how the right to jury might make defendants worse off.\textsuperscript{14} First, it is unclear how the jury right would positively or negatively affect the 90% of defendants who never go to trials. Prosecutors might indeed cede some sentencing issues considering the high costs of jury trials. But some defendants might not bother invoking such a right to protect their interests.\textsuperscript{15} What’s more, prosecutors are powerful. Right to jury might incentivize prosecutors to charge defendants more aggressively and exercise various tactics to coerce defendants to plead guilty, which “threatens serious unfairness.”\textsuperscript{16} Justice Scalia commented that, in arguing enforcing the element rule being unfair to defendants, the only authority Breyer cited is “an article written not by a criminal defense lawyer but by a law professor and former prosecutor.”\textsuperscript{17}

\textbf{II.3. Beneficial to Defendants}

King & Klein dismissed Bibas’ argument as “indeed startling” and “also dead wrong.” King & Klein argued that the element rule did not strengthen prosecutors’ bargaining positions by offering them additional bargaining chips. Before \textit{Apprendi}, prosecutors could force defendants to surrender to sentence enhancement facts by exercising the same coercive power, such as the threat of increasing sentence with a prior conviction. The element rule neither strengthened nor weakened such power.\textsuperscript{18}

On the contrary, \textit{Apprendi} empowered criminal defendants because it imposed a substantial procedural barrier that made it more difficult for prosecutors to punish defendants with sentence enhancements. \textit{Apprendi} raised burdens of proof and “makes it more difficult for prosecutors to prove aggravating facts that trigger longer sentences.” The rule also makes defendants less likely

\textsuperscript{13} 124 S. Ct. 2531
\textsuperscript{14} \textit{Id}. 2553
\textsuperscript{15} \textit{Id}. 2555
\textsuperscript{16} \textit{Id}. 2553
\textsuperscript{17} \textit{Id}. 2542
\textsuperscript{18} Stephanos Bibas, \textit{Apprendi} and the Dynamics of Guilty Plea. Stanford Law Review, Vol 54 No. 2 (2011) 311, P 311-312 (Bibas replied to King & Klein’s critiques, argued that King & Klein missed the fact that “\textit{Apprendi} has changed the worth of these bargaining chips.” Prosecutors can use the strong bargaining power to pressure defendants to give up enhancement issues.)
to admit to aggravating facts because they can threaten prosecutors with jury trials. Jury trials are expensive and rational prosecutors should try to avoid them.

In the majority opinion of Blakely, Justice Scalia cited the article by King & Klein, dismissing Justice Breyer’s concern that the right to jury is unfair to defendants for possibly increasing prosecutors’ bargaining power. Scalia reasoned that plea bargain was existent when sentence enhancement facts were considered as “sentencing factors.” Compared with bargaining over “sentencing factors,” bargaining over “elements” is more likely to yield favorable outcomes to criminal defendants because it is more effective to threaten prosecutors to prove crime elements.

What’s more, in the broader historic context, the jury functions as “the democratic branch of the judiciary power – more necessary than representatives in the legislature” and performing an important institutional check on the government. What’s more, it is a general view that right to jury operates as an important “structural democratic constraint” on legislative and executive branches in shaping criminal laws and procedure.

Legislatures play a dominant role in defining the criminal justice institution. The prominent example is the enactment of mandatory sentencing guidelines, at both state and federal levels, which dictates the procedure and outcomes of defendants’ sentencing. Despite the inherent democratic legitimacy of legislative branches, the political environment of being “tough on crime” might incentivize legislators to distort criminal laws by imposing excessive punishments on defendants and undermining the constitutional due process protections. Professor William Stuntz pointed out that legislatures are incentivized to empower police and prosecutors because of the potential gains from strong law enforcement. Scholars advocated more active judicial review to

---

19 Blakely v. Washington, 124 S.Ct. 2531, 2541
20 Id. 2542
21 Barkow, Recharging the Jury, Footnote 109
22 Legislators often express support or even propose legislative reform that lead to harsher criminal punishments. For example, senator Scott Cyrway sponsored a bill which raise the maximum penalty for importing heroin to Maine from 5 to 10 years. See Lawmakers consider tougher penalties for bringing heroin into Maine. https://www.pressherald.com/2016/01/25/lawmakers-consider-tougher-penalties-for-bringing-heroin-into-maine/ Another example, Senator Ted Cruz proposed a law which would punish undocumented immigrants with mandatory minimum of 5 years for reentering the US if had previously convicted of aggravated felony or illegally entering the US twice. Jordan Carney, Rubio backs Cruz on tougher penalties for illegal immigrants. The Hill https://thehill.com/blogs/floor-action/senate/257867-rubio-backs-cruz-on-tougher-penalties-for-illegal-immigrants
ensure robust due process protections, including jury trial protection, against legislatures’ manipulation of substantive criminal laws.\textsuperscript{24}

The right to jury can also help to mitigate the risk of arbitrary plea bargain decisions from prosecutors. In plea bargains, the jury serves as a particularly important institutional constraint as judicial review of plea deals is equivalent to cursory rubber stamp.\textsuperscript{25} Rachel Barkow criticized the use of mandatory sentencing guidelines for undercutting the benefits of jury trial options as a threat to prosecution.\textsuperscript{26} The mandatory guidelines authorized judges to enhance defendants’ sentences up to the statutory maximum by finding “relevant conduct” by a mere preponderance of evidence standard. As a result, it is still possible that defendants can be subjected to sentence enhancement based on facts that juries explicitly acquitted.

II.4. Consolidating the Competing Views with a Conceptual Framework

I develop a simple conceptual framework that accommodates the competing views between Bibas and King & Klein on how the element rule affects criminal defendants. Under the model, prosecutors decide to seek sentence enhancements against criminal defendants who committed the crime with sufficient severity. I assume that prosecutors’ underlying objective is to maximize punishment.\textsuperscript{27} Sentence enhancements increase defendants’ original sentence lengths. Therefore,

\begin{itemize}
\item \textsuperscript{24} Nancy J. King & Susan R. Klein, Essential Elements, 54 Vand. L. Rev. 1467 (2001), P1498 (Proposed a multi-factor balancing test to review the constitutionality of substantive criminal law statutes, which “preserve the legislative legitimacy in the defining crime and punishment, without wholesale abandonment of judicial control over the scope and application of constitutional criminal procedural guarantees.”)
\item \textsuperscript{25} Gerard E. Lynch, Our Administrative System of Criminal Justice, Fordham Law Review P2122 (Judges do not have enough information to make intelligible determination on defendants’ guilt. Therefore, in reviewing plea deals, judges only make sure if defendants are of sound minds and understand the consequences of accepting pleas.)
\item \textsuperscript{26} George Fisher, Plea Bargaining’s Triumph, 109 Yale L.J. 857 (2000) [Need to read 1039]
\item \textsuperscript{27} Rachel E. Barkow, Recharging the Jury: The Criminal Jury’s Constitutional Role in an Era of Mandatory Sentencing, 152 U Pa. L. Rev. 33 (2003). P96
\end{itemize}
punishment-maximizing prosecutors will maximize the charge of sentence enhancements under the constraint of evidentiary strength.

Let random variable $X$ represent the severity of how a defendant committed a given crime. $X$ does not capture the severity differentials across different types of crimes. Say, the difference between homicide and assault. Rather, $X$ speaks to how severe a given type of crime was committed. For example, defendants who committed aggravated assault are assigned with higher severity score $X$ if their cases are associated with one or a combination of commonly seen aggravating factors, such as cruelty, hate crime, and committing crimes against minors. Severity $X$ is also related to the evidentiary strength of relevant aggravating factors. A defendant is associated with higher severity if the racially biased motive against the victim is inferred from a Facebook post or tweet, rather than hearsay evidence.

A criminal defendant with severity level $X$ merits a sentence enhancement if $X$ is above a cut-off value, $C$. The bell curve in Figure 3.1 represents the distribution of severity $X$. $C_1$ represents the cut-off value that determines defendants’ take-up of sentence enhancements before Apprendi. Under the model, prosecutors pursue a sentence enhancement against a defendant with severity $X$ larger than $C_1$. Enforcing the element rule shifts the threshold value to the right, from $C_1$ to $C_2$. Prosecutors can only seek sentence enhancements against defendants with severity larger than $C_2$. Therefore, the new threshold decreases the fraction of criminal defendants punishable by sentence enhancements.\(^{28}\)

---

\(^{28}\) University Press (P9 Argued that prosecutors prefer longer sentences which give them additional leverage to solicit guilty plea and to avoid costly trials.)

\(^{28}\) I can show this proposition more rigorously. Suppose the cumulative distribution function of $X$ is $F()$. If the cut-off value is $C_1$, the share of defendants punishable by sentence enhancements is $1 - F(C_1)$. Under the cut-off value $C_2$, the share of defendants punishable by sentence enhancements is $1 - F(C_2)$. It is straightforward to show that $1 - F(x)$ is decreasing in $x$. Since $C_1 < C_2$, then $1 - F(C_1) > 1 - F(C_2)$. Therefore, a higher cut-off value decreases the share of defendants punishable by sentence enhancements.
Enforcing the element rule also represents a new constraint on prosecutorial discretion to punish defendants with sentence enhancements. Legal scholars characterize discretion as the power of public officials to make free choices within pre-set legal boundaries. 29 “Preponderance of evidence” is a low evidentiary standard, which gives rise to broad prosecutorial discretion to punish defendants with sentence enhancements. In contrast, the element rule requires prosecutors to prove sentence enhancement facts beyond a reasonable doubt, which substantially limits their free choice to pursue sentence enhancements against criminal defendants. 30 A higher burden of proof standard sets a more restrictive legal boundary and shrinks the population of defendants punishable by sentence enhancements. If prosecutors step out the legal boundary, defense attorneys are likely to demand prosecutors to prove relevant facts beyond a reasonable doubt, which might be too costly to attain, if not unattainable.

I use the term, “direct effects,” to define the effects of enforcing the element rule on prosecutors’ decisions to pursue sentence enhancements. The model predicts the direct effects to be negative. My simple conjecture is consistent with the theory from King and Klein that the element rule benefits criminal defendants by raising the procedural “price” of proving enhancement facts and empowering defendants to contest the validity of those alleged facts. 31 At

---

29 Marc L. Miller & Ronald F. Wright, The Black Box, 94 IWOA L. REV. 125, 167 (2008) (“This concept of discretion looks for enforceable legal standards that executive-branch actors must follow.”) Webster v. Doe, 486 U.S. 592, 592 (Held that when a statute is drawn in broad terms, judiciary is unable to review agency’s exercise of discretion due to a lack of meaningful legal standard.)

30 Andy Yuan & Spencer Cooper, Racial, Gender Disparities and State Prosecutors’ Discretion: Evidence from Blakely v. Washington, working paper, 13 (Develop a simple model showing that a higher burden of proof affects prosecutors’ choice of pursuing sentence enhancement by decreasing the expected benefits and increasing the overall costs.)

31 King & Klein, supra note X, 301
the same time, if prosecutors can drive positive utility/benefits from punishing defendants with sentence enhancements, the element rule also diminished the expected benefits for prosecutors to pursue sentence enhancements. Compared with judicial finding, juries are much less likely to find aggravating factors in applying the more stringent beyond a reasonable doubt standard.

However, this is not the end of the story. An additional constraint on prosecutorial discretion to charge sentence enhancements might result in prosecutors’ behavioral changes in making other prosecution-related decisions. In the US criminal justice system, prosecutors exercise broad discretion to determine how to process criminal cases, including what charges to bring, whether to prosecute, whether to seek pre-trial detention, what to convict, and what sentence to pursue. Consider the following facts from a hypothetical criminal case: an intoxicated man yelled obscenities, brandished a knife, and asked people for money in a subway station. He grabbed the arms of a woman who did not give him money. Then, the police arrested him. This set of simple facts enable prosecutors to charge defendants with any combination of 8 plausible offense, ranging from disorderly conduct to aggravated assault. Prosecutors are also free to press no charges or to defer prosecution. These decisions are essentially unreviewable.

To compensate the loss of opportunities to punish defendants with sentence enhancements, prosecutors may simply change how they exercise discretion and punish defendants through alternative means. I define such prosecutors’ compensating behavior as “spillover effects,” meaning that the effects of enforcing the element rule on prosecutors’ decisions that are related to sentence enhancements. The notion of “spillover effects” lies at the core of Bibas’s provocative claim that enforcing the element rule is detrimental to defendants’ interests. It is within the legitimate exercise of prosecutorial discretion to drive harder plea bargains if prosecutors are unwilling to see defendants go free from sentence enhancements.

In sum, I consolidate the competing views between Bibas and King & Klein using a simple dichotomy of “direct effects” and “Spillover Effects.” First, consistent with King & Klein, the

---

32 Christopher Robertson, Shima Baradaran Baughman & Megan S. Wright, Race and Class: A Randomized Experiment with Prosecutors, 16 Journal of Empirical Legal Studies 807, 820-821 (The authors ran a survey experiment and recruited actual prosecutors as experimental subjects. In the experiment, prosecutors were asked to make a wide range of charging decisions based on this particular set of facts.)
33 Id. 821 (The 8 plausible charges include disorderly conduct, loitering, public nuisance, criminal nuisance, harassment, endangerment, assault, and criminal assault.)
34 Id.
35 McCleskey v. Kemp, 481 U.S. 279, 280 (Courts presume that prosecutors’ exercise of discretion is reasonable. Overcoming such presumption requires “clear proof.”)
36 Bibas, Supra note __, at 1160.
element rule has direct “selection effects” which implies that prosecutors respond to the increase of procedural “price” by charging less sentence enhancement. Second, the element rule also has indirect “spillover” effects which manifest prosecutors’ compensating behavior in treating defendants more harshly in offering plea deals, in line with Bibas’s argument. Ultimately, whether the right to jury is harmful to the defendant is an empirical question.

II.5. Enforcing Right to Jury and Unwarranted Outcome Disparities

Congress passed the Sentencing Reform Act in 1984, replacing the previous indeterminate sentencing system with a determinate one based on the mandatory sentencing guidelines. At state level, many states adopted sentencing guidelines which either require or encourage judges to follow when making sentencing decisions.\(^{37}\) An important goal of adopting sentencing guidelines is to reduce the warranted disparities among similarly situated defendants.\(^{38}\) Invalidating mandatory sentencing guidelines based on Sixth Amendment will inevitably affect outcome disparities.

Legal academics and professionals predominately focus on how enforcing right to jury affects sentencing disparities by changing judicial discretion. In United States v. Booker, justices debated the appropriate remedy for declaring the federal mandatory sentencing guidelines unconstitutional, either making the Guidelines advisory or engraving jury factfinding requirement onto the Guidelines. Justice Breyer wrote the majority opinion which adopted the former approach.\(^{39}\) An important reason why Breyer opposed the latter solution is that it will undermine the legislative goal of ensuring defendants with similar conduct receive similar sentences.\(^{40}\) If the

\(^{37}\) State Sentencing Guidelines, Profiles and Continuum, Page 4 (Report from the National Center for State Courts) (Link: https://www.ncsc.org/__data/assets/pdf_file/0022/25474/state_s...guidelines.pdf) (The report surveyed 21 states that adopted sentencing guidelines. The report evaluated the nature of state sentencing guidelines based on the following 6 criteria: (1) is there an enforceable rule related to guideline use; (2) is the completion of a worksheet or structured scoring form required; (3) whether a sentencing commission regularly report on guideline compliance; (4) whether compelling and substantial reasons required for departure; (5) whether written reasons are required for departure; and (6) whether there is appellate review of defendant-based challenges related to sentencing guidelines. State sentencing guidelines are rated on a scale from 1 to 12 with 1 being the most voluntary and 12 being the most mandatory.)

\(^{38}\) U.S.C. § 3553(a)(6) “The Need to avoid unwarranted sentence disparities among defendants with similar records who have been found guilty of similar conduct” Empirical research also demonstrated that states’ adoption of sentencing guidelines, either voluntary or mandatory, is associated a decrease in defendants’ sentencing disparities. Infra Note X (Cite Pfaff’s work)

\(^{39}\) United States v. Booker, 543 U.S. 220, at 246

\(^{40}\) Id, at 250
law prohibits judges from considering relevant sentencing facts that juries did not find by beyond a reasonable doubt, defendants’ sentence lengths are unlikely to reflect their “real conducts,” which will exacerbate sentencing disparity.

On the flip side, allowing substantial judicial discretion in determining “real conducts” might be a new source of sentencing disparities. As Justice Stevens pointed out, there is a lack of effective legal mechanisms to ensure that judges do not abuse their discretion in determining defendants’ “real conducts.”41 On the one hand, the laws are unclear to what extent defendants’ “real conducts” matter in determining their sentence lengths, which left entirely to judges’ discretion.42 There is also lack of meaningful appellate review standard on judges’ determination of “real conducts.”43

There are relatively fewer discussions on impacts of enforcing right to jury on prosecutorial discretion to seek sentence enhancements, which also affect outcome disparities among defendants. In many states, prosecutors, instead of judges, make the de facto decisions on whether defendants receive sentence enhancements. For example, in North Carolina, judges are unable to enhance defendants’ sentence lengths unless prosecutor requests the enhancements and proves the relevant aggravating facts.44 In these jurisdictions, requiring prosecutors to prove relevant facts by beyond a reasonable doubt will inevitably impose a significant constraint on their discretion to pursue sentence enhancements.

III. The Court’s Post-Apprendi Enforcement of Jury Trial Guarantee

The Apprendi rule triggered a series of Supreme Court’s decisions which more rigorously enforced the Sixth Amendment jury trial guarantee and fundamentally transformed the pre-existing sentencing laws. First, the Court extended the scope of the element rule to include facts that lead to sentence beyond “mandatory minimum,” not just “statutory maximum.” What’s more, Apprendi also lead to Blakely v. Washington and Unites States v. Booker, which is directly and indirectly lead to the demise of mandatory sentencing guidelines at both state and federal level. This section reviews the key post-Apprendi development of jury guarantee.

41 Id, at 288
42 Id.
43 Id
44 Infra Note X
III.1. Mandatory Minimum

The Court went back and forth in extending the Apprendi rule to mandatory minimum sentences for federal crimes. Initially, the Court refused to extend jury guarantee to cover mandatory minimum in Harris v. United States.\(^{45}\) In Harris, the defendant William Harris was convicted of illegal possession of firearms in relation to drug trafficking. In sentencing, the judge subject the defendant to a mandatory minimum of 7 years, instead of the regular 5 years, based upon his own finding that the defendant “brandished” the weapon in committing crime, although “brandishing” weapon was neither indicted nor tried.\(^{46}\) The defendant challenged the sentencing decisions and argued that “brandishing” weapon is an element of crime, which only jury can find by beyond reasonable doubt.\(^{47}\)

The Court drew a distinction between facts that raise the statutory maximum and facts that increase mandatory minimum. It held that the latter are sentencing factors which judge can find by preponderance of evidence. The majority opinion declined to broadly apply the Apprendi reasoning and emphasized the “continuing vitality” of McMillan v. Pennsylvania in which the Court allowed judges’ discretionary factfinding as a legitimate basis of imposing mandatory minimum sentences.\(^{48}\) Harris marked as an important boundary against broad interpretation of the Apprendi jury guarantee.

In 2013, the Harris decision was formally distinguished by Alleyne v. United States in which the Court held that facts that increase defendants’ mandatory minimum sentence also trigger a jury process.\(^{49}\) The issue in Alleyne is exactly the same as the one in Harris, whether “brandishing” weapon under 18 U.S.C. § 924(c)(1) is a sentencing factor or element of crime. The Court abandoned the prior distinction between facts increasing the statutory maximum and facts increasing the mandatory minimum in determining whether jury process applies.\(^{50}\) It recognized that sentencing floors are as relevant as ceilings in defining criminal punishment.\(^{51}\) Therefore, facts

\(^{45}\) 536 U.S. 545 (2002)
\(^{46}\) Id. at 551 citing 18 U.S.C. § 924(c)(1)(A)
\(^{47}\) Id.
\(^{48}\) Id. at 556-557
\(^{49}\) 570 U.S. 99 (2013), at 99
\(^{50}\) Id, at 109
\(^{51}\) Id, at 112-113
that increase defendants’ mandatory minimum sentences also need to be included in the indictment and proved to a jury beyond a reasonable doubt. Justice Sotomayor wrote a separate concurrence opinion explaining why the facts in Alleyne merit a departure from the stare decisis principle.\textsuperscript{52} First, right to jury is a procedural rule, which generally does not implicate as strong reliance interests as substantive rules.\textsuperscript{53} Moreover, Harris has minimal precedential values in light of the Court’s previous decisions that struck the mandatory sentencing system at both state and federal levels.\textsuperscript{54}

The Alleyne decision facilitated the criminal justice reform during the Obama administration.\textsuperscript{55} Immediately following the Alleyne decision, Attorney General Eric Holder issued a sweeping memorandum instructing federal prosecutors to suspend charging mandatory minimum sentences among defendants associated with “low-level and non-violent drug offenses.”\textsuperscript{56} However, later Attorney General Jeff Sessions suspended the policy four months into the Trump administration.\textsuperscript{57} Despite such policy shift, the Alleyne rule still constrain the government from arbitrarily charging defendants with mandatory minimum sentences as it requires the federal prosecutors to prove relevant facts to a jury beyond a reasonable doubt.

III.2. \textit{Blakely v. Washington} and the demise of mandatory sentencing guidelines

\textsuperscript{52} Id at 118 (Departure from Stare Decisis require a special justification beyond the argument that the precedent was wrongly decided.)
\textsuperscript{53} Id. at 119
\textsuperscript{54} Id. at 120
\textsuperscript{56} See Department Policy on Charging Mandatory Minimum Sentences and Recidivist Enhancements in Certain Drug Cases, Memorandum to the United States Attorneys and Assistant Attorney General for the Criminal Division (\url{https://www.justice.gov/sites/default/files/oip/legacy/2014/07/23/ag-memo-department-policypon-charging-mandatory-minimum-sentences-recidivist-enhancements-in-certain-drugcases.pdf}) P2 (The Memo instructed the federal prosecutors not to seek mandatory minimum sentences if defendants meet four prescribed criteria.); Also See Stephanie Holmes Didwania, Mandatory Minimums and Federal Sentences (Working paper) P4 (This is the first study evaluating the effectiveness of Holders’ Memo on defendants outcomes. It found that eligible defendants are 40\% less likely to be charged with mandatory minimum sentences. But their sentence lengths remained almost unchanged.)
\textsuperscript{57} Memorandum for All Federal Prosecutors, Department Charging and Sentencing Policy issued on May 10\textsuperscript{th}, 2017 See the link: \url{https://www.justice.gov/archives/opa/press-release/file/965896/download} (The memo instructed the federal prosecutors to pursue the most serious offenses that carry the most substantial guideline sentences, including the mandatory minimum sentences.)
In *Blakely*, the defendant was initially charged with first-degree kidnapping for abducting his wife. Through a plea agreement, he was convicted of a lower-level offense, second-degree kidnapping in relation to domestic violence and use of firearms, which translates to a Class B felony under the state law.\(^{58}\) In Washington, judges used to determine defendants’ sentence lengths according to a mandatory sentencing guideline, which prescribe standard sentencing ranges on the basis of the class of convicted offense and prior convictions. The state law allowed judges to sentence defendants above the standard range if they find “substantial and compelling” reasons. The law prescribed an illustrative list of aggravating factors as a basis for such upward departures.

The state prosecutors initially sought sentences from the “standard range” between 49 months and 53 months. However, judges issued an exceptional sentence of 90 months, 37 months longer than the maximum sentences of prosecutors’ recommendation but shorter than the maximum imprisonment period for Class B offense, 10 years. The judge justified the upward departure based on his finding that the defendant acted with “deliberate cruelty,” a statutorily prescribed aggravating factor.

*Blakely* uniquely involved two sentence ceilings, the maximum sentence of the standard range applied to the defendant, 53 months, and the maximum imprisonment for Class B felony, 10 years. The main issue is which sentence ceiling is the “statutory maximum.” The state argued that defendant’s sentence length did not trigger the *Apprendi* rule because judges’ finding of “deliberate cruelty” did not lead to a sentence beyond the maximum sentence for Class B felony. In contrast, defendant argued the statutory maximum is the upper end of the applicable standard range. In a 5-4 decision, the Court ruled in favor of the criminal defendant and struck down the entire mandatory sentencing guideline in Washington.

*Blakely* is perhaps the most consequential case because it struck down the entire mandatory sentencing guidelines in Washington, which had been in place since the state legislature’s passage of the Sentencing Reform Act of 1981. The decision also jeopardized the constitutionality of the mandatory sentencing guidelines at state and federal levels. Sixth months after the *Blakely* decision, the Court decided the *United States v. Booker*, which struck down the federal mandatory sentencing guideline for violating defendant’s Sixth Amendment and declared the guidelines effective advisory.\(^{59}\)

\(^{58}\) §§9A.40.030(1), 10.99.020(3)(p), 9.94A.125

\(^{59}\) 543 U.S. 220, at 221-222 (2005)
After Booker, the Court issued additional decisions clarified the appellate review standard of judges sentencing decisions, which further enhanced judges’ sentencing discretion. In Rita v. United States, the Court held sentences from the Guidelines range are presumptively “reasonable” as they appropriately reflect the judgment of the Federal Sentencing Commission.\textsuperscript{60} In Gall v. United States, the Court held that downward departures from the regular Guidelines range are not presumptively unreasonable and do not require “extraordinary” circumstances.\textsuperscript{61} Following the Court’s ruling in Gall, appellate review of sentences should follow the deferential abuse-of-discretion standard in reviewing sentences outside of the Guidelines range.\textsuperscript{62} In Kimbrough v. United States, the Supreme Court held that the district judge did not abuse his discretion when sentenced the defendant 4.5 years below the Guidelines range based on a policy disagreement with the crack/powder sentencing disparity in the Guidelines.\textsuperscript{63}

At the state level, Blakely triggered a wave of state legislative efforts to fix their constitutionally flawed sentencing guidelines. Some states injected the requirement of jury trials in the procedure of imposing sentencing enhancement on criminal defendants. Some states followed the Booker approach and converted the previously mandatory guidelines advisory. Some states maintained the binding guidelines.

IV. Review of Relevant Economics Literature

This section reviews the two main strands of empirical literature this paper contributes to. First, there is an empirical literature related to the right to jury. Second, there is an economic literature on prosecutors’ exercise of discretion, both theoretical and empirical. My empirical analysis bridges these two relatively distinctive strands literature by demonstrating, theoretically and empirically, the direct and indirect effects of right to jury on the exercise of prosecutorial discretion.

IV.1. Empirical Literature on Right to Jury

\textsuperscript{60} 551 U.S. 338, at 347 (2007)
\textsuperscript{61} 552 U.S. 38, at 47 (2007)
\textsuperscript{62} Id. at 60 - 61
\textsuperscript{63} 552 U.S. 85, at 111 (2007); at 95-96 (The crack/powder sentence disparity originates from the Anti-Drug Abuse Act of 1986. In determining defendants’ minimum sentences, the law “treats every gram of crack cocaine as the equivalent of 100 grams of powder cocaine.”)
There is a set of empirical papers examining how expanding judicial discretion affects judges’ sentencing behavior. This literature evaluates how Booker affects various aspects of sentencing outcomes. In Booker, the Court struck down federal mandatory sentencing guidelines violating defendants’ Sixth Amendment right to jury, which expands judicial discretion in sentencing. The Booker decision creates a unique natural experiment as the federal sentencing guideline no longer constrained judges’ sentencing decisions. Researchers attempt to empirically detect whether Booker yielded behavioral changes from federal judges.

One set of papers focuses on examining changes of inter-judge disparity resulting from expanding judicial discretion. Existing studies show that Booker increased the inter-judge sentencing disparity federal courts. In other words, Booker made “harsh” judges harsher and “lenient” judges more lenient. For example, Professor Crystal Yang found that compared with the average judges, the “extra” sentencing length associated with a one-standard-deviation “harsh” judges increase from 2.8 months, the pre-Booker level, to 5.9 months, the post Kimbrough/Gall level. In comparison, chance of receiving a “below-range” departure associated with a one-standard-deviation “lenient” judges increased from the pre-Booker 4.7% to the post-Kimbrough/Gall 6.9%. What’s more, judges’ sentencing discretion also varies at state level as some states adopted mandatory sentencing guidelines while others chose advisory ones. Professor John Pfaff estimated how the adoption of the two different types of guidelines differentially affect sentencing disparity. In theory, mandatory guidelines should make defendants’ sentencing lengths more cohesive as they put more stringent restrictions on judges’ exercise of discretion

---

64 See Ryan W. Scott, Inter-Judge Sentencing Disparity after Booker: A First Look, 63 Stan. L. Rev. 1 (2010). The Study by Scott is the first empirical investigation on the inter-judge disparity. The limitation of the study is that the dataset only covers cases decided by around 10 federal judges from the District of Massachusetts. P40-41. Examining changes of R-squared in linear regressions, Scott finds that Kimbrough/Gall lead to a “doubling” effects of judge assignment on sentence length. In addition, the inter-judge sentencing disparity in sentence length is larger than 2 years in cases which are not subject to mandatory minimum. See Crystal S. Yang, Have Interjudge Sentencing Disparity Increased in an Advisory Guideline Regime – Evidence from Booker, 89 N.Y.U. L. Rev. 1268 (2014). P1293-1294, Yang studied the same question but assembled a comprehensive national-level, multi-district dataset which, for the first time, matched sentencing data with judge identifiers.

65 Also See Yang P1340-1341 (Provides a description of the analysis of variance statistical methods in estimating the inter-judge disparity in sentencing outcomes. The estimate of \( \sigma_y \) means the change of defendants’ sentencing outcomes, such as sentencing length or the below-range departure, associated with a one-standard-deviation increase of judges’ “harshness.”)

compared the advisory counterparts.\textsuperscript{67} Indeed, Pfaff finds that, adopting mandatory sentencing guideline is associated with as large as 57\% decrease in the pre-guideline sentencing length variation, compared with a 35\% decrease among states adopting voluntary guidelines.\textsuperscript{68}

Another set of \textit{Booker}-studies focuses on how it changes racial sentencing gap. The results in this literature shows that \textit{Booker} did not immediately change the racial sentencing gap in the short run but increased the gap by a sizable magnitude in the long run. Using a large and comprehensive dataset of federal criminal cases, Professors Sonja Starr and Marit Rehavi employs a regression discontinuity-style approach and examine the change of racial sentencing gap over a relatively short time window, 12 and 18 months before and after the \textit{Booker} decision.\textsuperscript{69} The main finding does not suggest that switching to an advisory sentencing regime lead to a big increase in racial sentencing gap in the short run.\textsuperscript{70} In contrast, Crystal Yang investigated the relatively long-run effects of \textit{Booker} on racial sentencing gap over an extended sample period from 1994 to 2010.\textsuperscript{71} The results suggest that \textit{Booker} widened the black-white sentencing gap by 2 month, which doubled the pre-\textit{Booker} baseline gap.\textsuperscript{72} In addition, Yang further showed that the increasing black-white sentencing disparity is largely attributable to the set of post-\textit{Booker} judges who tend to impose longer sentences on black defendants than the pre-\textit{Booker} counterparts.\textsuperscript{73}

\textsuperscript{67}Michael Tonry, Sentencing Matter, Oxford University Press (p28-29, argued that voluntary guidelines are ineffective in regulating judges’ sentencing behavior, claiming that “evaluation showed that voluntary guidelines typically had little or no demonstrable effect on sentences imposed… and in most places they were abandoned,” with possible exception of the voluntary guideline in Delaware.

\textsuperscript{68}Supra Note X Pfaff UCLA L. Rev. P239

\textsuperscript{69}Sonja B. Starr; M. Marit Rehavi, Mandatory Sentencing and Racial Disparity: Assessing the Role of Prosecutors and the Effects of Booker, 123 Yale. L. J. 2 (2013), 52 (Explained their empirical strategy by saying “we … look for sharp breaks in these trends – discontinuities – immediately after \textit{Booker}. This approach is … a regression discontinuity-style estimator.”) See Sonja Starr, Did Booker Increase Disparity? Why the Evidence Is Unpersuasive, Federal Sentencing Reporter, Vol. 25 No.5 (2013), 323, 323 (Criticized the empirical approach of Federal Sentencing Commission in evaluating the effects of \textit{Booker} on racial sentencing gap.)

\textsuperscript{70}Starr & Rehavi YLJ P64-65 (They plot the residualized racial sentence gap which appears to be relatively stable across all time periods. The regression tables specified the time window for each specification. Also, P60, the estimated regression coefficients for black-white difference in the discontinuity are not statistically significant.)


\textsuperscript{72}Yang JLS (2015) P77 (Also find the direct mechanisms for the increase in sentence disparity: “Increased racial disparities in sentence length can be attributed to black defendants being more likely to be sentenced above the guidelines-recommended range and less likely to be sentenced below the guidelines-recommended range, compared with similar white offenders.”)

\textsuperscript{73}Id P102-104 (Yang relies on random assignment of cases to identify the difference in sentencing behavior between pre-\textit{Booker} and post-\textit{Booker} judges. The results suggest that post-\textit{Booker} judges sentence black defendants 5.5 months longer than white defendants. Furthermore, judges’ tenure is not predictive of the sentencing disparity. The inter-judge sentencing difference is primarily attributable to the exposure to the previous mandatory sentencing guidelines, as opposed to judges’ experience.)
There exists only a small body of work examining how the development of jury protection affects sentencing outcomes of offenders convicted of state crimes, which constitutes more than 90% of the criminal cases in the US. The existing state crime studies are motivated by the Blakely decision, which explicitly invalidated the mandatory sentencing guideline in Washington for violating defendants’ Sixth Amendment Right to juries.

There is a body of descriptive work on how states complied with Blakely by amending the laws. Empirical work on how Blakely affects state criminal justice system remains to be scarce. Iannacchione and Ball made the first step to examine how Blakely affected upward departure in defendants’ sentence length by analyzing the Washington state crime data. They find that Blakely lead to a decreasing likelihood of upward departure by around 70%. The results are consistent with my findings but of larger magnitude. However, their study has limitations. Empirically, the analysis was based on a relatively small sample which lacks key demographic variables, such as location, crime type and attorney representation. Theoretically, the authors could not attribute the observed decrease in upward departure to prosecutors’ behavioral changes resulting from higher burdens of proof in proving relevant facts. This limitation results from the old sentencing practice in Washington state where judges’ upward departures in sentencing decisions did not require prosecutors’ proof of aggravating factors at all. Such institutional feature made it very difficult to link sentencing upward departure to prosecutorial discretion. I am able to largely address these limitations using a comprehensive administrative dataset, with rich demographic characteristics, studying the impacts of Blakley in North Carolina sentencing where judges are never free to make upward departures without prosecutors’ proof of relevant facts.

IV.2. Literature on Prosecutors’ Behavior

74 Cite
76 Id. P429 (Refer to the main results in Table 3.3. “In fact, the odds of a case receiving an upward departure after the Blakely decision are 69% less than the odds of receiving an upward departure before the Blakely decision.”)
77 Id. P432 (“this study is a product of using secondary data … there is an absence of relevant variables in the model.”)
78 § 9.94A.120(2) & (3) (judges can unilaterally increase sentences if they find substantial and compelling reasons so long as they provide written opinions on the departure.)
There is an economic literature which examines prosecutors’ behavior both theoretically and empirically. Theoretical work documents various models on how prosecutors decide between plea bargain and trial under different assumptions. William Landes developed the first economics-based analysis on how prosecutors make decisions by maximizing expected punishment subject to limited resources. Landes’s model justifies plea bargain for saving material inputs in litigation. Professors Gene Grossman and Michael Katz proposed an alternative framework where prosecutors’ objective is to maximize social welfare. The model assumes private information on defendants and favors plea bargain as effective insurance and screening device. Jennifer Reinganum further extended the Grossman-Katz model by developing a game theoretical framework which made the following two advancements: (1) assume correlated private information on both prosecution and defense; (2) characterizing prosecutors’ optimal offers under both complete and restricted discretion.

My simple economic framework enriches the literature by characterizing how “procedural prices” affect prosecutors’ decision to charge sentence enhancement.

Given the enormous and largely unchecked prosecutorial discretion, empirical analysis of prosecutors’ behavior yields valuable insights for institutional reform which constrains prosecutors’ exercise of power. One set of empirical literature uses data to estimate prosecutors’ behavioral responses various institutional changes. Yang examined how prosecutorial decisions respond to judicial vacancies which realistically affect the costs of prosecution because vacancies increase judges’ caseloads, case duration and the likelihood of violating defendants’ right to speedy trials. If prosecutors are constrained by limited resources, judicial vacancies are likely to induce changes in prosecutors’ decisions on how to proceed criminal cases at the margin. The results suggest that greater scarcity of judges lead to two key behavioral change: (1) screening out cases with relatively weaker evidentiary basis by declining prosecution or dismissed filed charges and (2) offering more

79 Supra note Landes.
81 Jennifer F. Reinganum, Plea Bargain and Prosecutorial Discretion, American Economics Reivew, (1988) P715(The model implies that prosecutors can make individualized offers under complete discretion condition, while restrictive discretion only allows prosecutors to make one single sentence offer to all defendants. Also, defendant’s private information is whether he or she is guilty or innocent. Prosecutors’ private information is defendants’ likelihoods of conviction.)
favorable plea deals in exchange for guilty pleas. In addition, the literature also documents prosecutors’ behavioral changes in response to mandatory minimum laws.

Professors Daniel Kessler and Anne Piehl empirically evaluated the change of defendants’ sentencing outcomes resulting from the passage of California’s Proposition 8, which increased sentence duration for repeated offenders convicted of “qualifying” crimes. The results show that there was a 50% increase in sentencing length among repeated offenders convicted of crimes covered by the law. The law also had spillover effects where sentencing length also increased, small but statistically significant, among repeated offenders not covered by Proposition 8. The authors argued that mandatory minimum laws give prosecutors additional bargaining power, which lead to an overall increase in the harshness of prosecution. A later study suggests that prosecutors responded to three-strike laws by reducing felony charges to misdemeanors among targeted offenders, which resulted from prosecutors’ exercise of discretion and reflected their preferences on who should be punished by the law. There does not exist any work related to how prosecutors change behavior in relation to changing burdens of proof.

Another line of work assesses the relationship between prosecutors’ exercise of discretion and outcome disparities, primarily on the basis of race and gender. Among offenders convicted of federal crimes, there exists substantial racial and gender sentencing gaps unfavorable to black and male defendants. Using a comprehensive administrative dataset, Rehavi and Starr showed that prosecutors’ discretionary decisions can explain a substantial fraction of racial and gender sentencing gaps. It is worth noting that studies by Rehavi and Starr demonstrate that prosecutors’ discretionary decisions introduced racial and gender disparities. However, they did not definitively

---

83 Id, P291 (Yang developed causal estimates by implementing instrumental variable strategy to correct the endogeneity bias in judges’ decisions to exit. The results suggest that 10% increase in judicial vacancies lead to an 4.17% increase in the probability of declining prosecution, 3.14% increase in the probability of dismissing filed charges. As to case outcomes, 10 percent increase in judges’ leniency increase the 0.38 percentage point increase in defendants’ likelihood of pleading guilty and 1.2 percentage point decrease in the probability of incarceration.)


86 M. Marit Rehavi & Sonja B. Starr, Racial Disparity in Federal Criminal Sentences, 122 Journal of Political Economy 1320, 1323 (2014) (Federal prosecutors’ decisions to bring charges with mandatory minimum can explain more than half the black-white sentence gap which unexplained by defendants’ pre-charge characteristics, such as arrest offense and criminal history). Sonja B. Starr, Estimating Gender Disparity in Federal Criminal Sentences, 12 American Law and Economics Review 127, 140-142 (2015) (In non-drug cases, guideline fact-finding explains around 60% of unexplained gender sentencing gap. In drug cases, minimum sentence charging decisions explains one third of the unexplained gap.)
prove prosecutors’ discriminatory intent because the disparities could be entirely driven by the unobserved characteristics associated with defendants’ race and gender. There is a rich criminology literature showing disparities among a wide range of prosecutorial decisions including initial screening, pretrial detention, charge dismissal, guilty plea and sentencing. The criminology literature uses selection-on-observable approach and largely ignored the omitted variable bias, which is the biggest empirical challenge of convincingly proving prosecutors’ discriminatory intents. CarlyWill Sloan exploits randomly assignment of misdemeanor cases to prosecutors and finds evidence suggesting prosecutors’ systematic racial bias against people of opposite race. Specially, being randomly assigned to prosecutors of opposite race lead to 8% increase in defendants’ probability of being convicted of property crimes. Through a similar approach, Stephanie Didwania shows federal prosecutors gender bias against defendants of opposite sex.

V. Sentence Enhancement in North Carolina

This section briefly reviews the North Carolina state sentencing law, which provides the context of the empirical analysis.

V.1. Structured Sentencing Guidelines

In the 1990s, North Carolina implemented sweeping criminal justice reforms to combat with widespread disparities and to concentrate prison resources on violent crimes. The key milestones include the passage of the Structured Sentence Act and the subsequent adoption of

---

87 Reference
88 CarlyWill Sloan, Racial Bias by Prosecutors: Evidence from Random Assignment, working Paper (2020). Also see Cody Tuttle, Racial Disparity in Federal Sentencing: Evidence from Drug Mandatory Minimum, working paper (2020) (The analysis shows the 2010 Fair Sentencing Act lead to a bunching of cases charged with trafficking 280 grams of crack-cocaine which leads mandatory minimum sentence. The results showed that the bunching is disproportionately driven by cases against black or Hispanic defendants).
89 Stephanie Holmes Didwania, Gender-Based Favoritism Among Criminal Prosecutors. Working Paper (2020)
structured sentencing Guidelines. The state law mandate judges to follow the Guidelines when making sentencing decisions.

Figure 3.2 shows the sentencing Guidelines that was effective during the sample period of this study, from 2001 to 2007. The state mandatory sentencing guideline consists of grid cells based on the severity of convicted offense and defendants’ criminal history. North Carolina uses a point system to quantify defendants’ criminal history. The state criminal law also introduces different classes of felony offenses to measure the severity. Under the guideline, more severe convicted offenses and longer criminal history inevitably result in longer imprisonment time.

Within each grid cell, the guideline specifies three different minimum sentence ranges, presumptive, mitigated, and aggravated. Judges choose a minimum sentence length from the presumptive range by default. In addition, for some grid cells, there are options to punish defendants with either incarceration or probation sentences. Judges will choose the type of punishment for cases that went to trial. Otherwise, prosecutors will choose the type of punishment over the course of plea negotiation with defendants.

V.2. Sentence Enhancement Procedure and Blakely v. Washington

Under the guideline, sentence enhancement is equivalent to an upward departure from the presumptive sentence range. The state law prescribes a list of aggravating factors, which lead to sentence enhancement.

To achieve an upward departure, the state needs to notify the Court of its intent to seek minimum sentences in the aggravated range. Under the current law, defendants either admit to aggravating factors or are entitled to juries trials to determine those factors. Proving some of the legally prescribed aggravated factors are difficult because they are subjective or involves

---

92 Id. at 2 (In 2009, the North Carolina legislature modified the guidelines by changing the minimum sentence lengths for Class B1 through G felonies. The change took effect December 2009.)
93 § 15A-1340.16. (d)
94 § 15A-1022.1. (a) Procedure in accepting admissions of the existence of aggravating factors in felonies
95 § 15A-1022.1. (b)
96 § 15A-1340.16. (d)(7)“The offense was especially heinous, atrocious, or cruel.”
defendants’ intent. Therefore, prosecutors’ decisions to charge aggravated sentences inevitably involve a great deal of judgment call or exercise of discretion.

The right to jury was never part of the North Carolina state criminal procedure in imposing sentence enhancement until the Blakely decision. It’s worth highlighting reasons why Apprendi did not result in sweeping skepticism over the constitutional legitimacy of the sentencing guideline. First, the Apprendi holding appears to be narrow and the meaning of “statutory maximum” was unclear. What’s more, in state v. Lucas, the North Carolina state Supreme Court explicitly rejected the interpretation that seeking minimum sentence in the aggravated range is equivalent to sentencing beyond “statutory maximum.” Additionally, the Court’s decision in Harris further constrained the broad reading of the Apprendi decision.

Prior to Blakely, judges determine the existence of aggravating factors by preponderance of evidence standard. Blakely made it crystal clear that sentence enhancement upon judicial factfinding is unconstitutional. In June 2005, the state legislature formally passed the “Blakely” bill, which amended the legal procedure of finding aggravating factors to comply with the Blakely decision. The amendment substituted the previous judicial factfinding rule with the new jury trial guarantee if defendants are charged with sentence enhancement but refuse to admit aggravated factors.

Blakely was decided in June 2004 and is likely to affect the State sentencing law practice earlier than the timing of the formal legislative amendment. Indeed, Ronald Wright documented that legal professionals from both prosecution and the criminal defense responded immediately to Blakely. On the one hand, prosecutors started dropping requests for seeking aggravated sentences as the high costs of going through jury trials do not outweigh the benefits of punishing defendants with

---

97 § 15A-1340.16. (d)(17)“The offense for which the defendant stands convicted was committed against a victim because of the victim's race, color, religion, nationality, or country of origin.”
98 548 S.E.2d 712 (N.C. 2001) (the state Supreme Court held that statutory maximum sentences refer to the maximum sentences set out in N.C.G.S. § 15A - 1340.17(e). Right to jury does not apply so long as defendants are not sentenced beyond statutorily prescribed maximum sentences which were determined by using the longest minimum sentence.) The interpretation of statutory maximum in Lucas was later overruled by State v. Allen, 615 S.E.2d 256 (N.C. 2005) (The state followed Blakely and held that “any fact that increase the penalty for a crime beyond the prescribed presumptive range, i.e., the ‘statutory maximum,’ must be submitted to a jury and proved beyond reasonable doubt.”)
99 See A Summary of Legislation in the 2005 General Assembly of Interest to North Carolina Public official, which consists of key provisions of the “Blakely” bill. The document can be found in the following link: https://www.sog.unc.edu/sites/www.sog.unc.edu/files/full_text_books/2005%20NC%20Legislation.pdf
100 Ronald Wright, Blakely and the Centralizers in North Carolina, Federal Sentencing Reporter (2005)
longer incarceration periods, typically a few more months. In addition, the criminal defense also started filing requests for jury trials if prosecutors sought sentence enhancement on their clients.

VI. Data and Empirical Strategy

VI.1. Data and Summary Statistics

The primary data source is the criminal case records from the Administrative Office of Court (AOC) in North Carolina. The data consists of all of the criminal cases disposed of at the state superior courts. I impose a number of restrictions to construct the final analysis sample. First, I drop all the non-felony charges because the sentencing Guidelines in Figure 3.2 only apply to felony. Second, I restrict the time frame of the disposition date between 2001 and 2007, 3 years before and after the Blakely decision. We supplement the main data source with county-level unemployment data from the Bureau of Labor Statistics (BLS) and arrest data from Uniform Crime Report (UCR). Unemployment data is a good proxy of local economic conditions. Arrest data captures police behavior.

The dataset has many attractive features for empirical analysis. First, it includes all of the criminal charges filed to the state court system and keeps track of case outcomes across various stages, including charging, plea bargain, and final disposition. Second, there is rich set of demographic characteristics about the defendants, such as name, gender, race, date of birth, criminal history, attorney representation, and offense types. The rich demographic information enables me to examine the heterogeneous effects of Blakely across different types of cases and defendant characteristics. I constructed a measure of sentence enhancement by mapping the observed minimum sentences to the corresponding guideline ranges. I infer the defendants to

---

101 Thanks to Bernardo Silveira who makes the data publicly available. See Bernardo Silveira, Bargaining with Asymmetric Information: An Empirical Study of Plea Negotiations, 85 Econometrica 419 (2017). (The data files and data cleaning code are available in the online supplemental material. The link: https://www.econometricsociety.org/publications/econometrica/2017/03/01/bargaining-asymmetric-information-empirical-study-plea)

102 I use the county identifiers to match the BLS and UCR data to the criminal case records.

103 I am able to find defendants’ sentencing grids by using the observed convicted offenses and criminal history points. Within the identified sentencing grid, I am able to determine if defendants’ minimum sentences fall in presumptive, aggravated, or mitigated ranges.
have received sentence enhancements if their observed minimum sentence lengths fall into the aggravated ranges.

Table 3.1 reports the key summary statistics of the analysis sample. There is a total number of 213,264 felony cases in the sample. 84% of the defendants are male. Defendants are Black or Hispanic in 58% of the cases. The average age of criminal defendants is around 30 years old. 24% of the defendants are represented by private attorneys and 22% by public defenders. The average number of criminal history points is about 4.6. The average arrest rate is about 536 arrests per 100,000 county population. The unemployment rate is 5.8.

As the case outcomes, 86% of the cases are resolved through plea bargains. The plea bargain rate increases to 95% after dropping cases that are voluntarily dismissed by prosecutors. 11% of the defendants plead guilty to lower-level felonies and 13% to misdemeanors. 31% of the defendants received incarceration punishment. The average minimum sentence length is around 17 months. 2% of the defendants received sentence enhancement where the minimum sentence length comes from aggravated ranges. 8% of the defendants received mitigated sentences.

Figure 3.3A shows the monthly average rate of cases with sentence enhancements. I calculated each data point using the raw data. The line in the middle is June 2004, the month of the Blakely decision. I fitted the scatterplot with two separate curves pre and post Blakely using the second-degree polynomials. The figure shows that Blakely caused a sharp discontinuity in the defendants’ likelihood of receiving sentence enhancements. The share dropped from roughly 3% to less than 1.5% initially. Figure 3.3B shows the residualized sentence enhancements. The Blakely-induced discontinuity remains to be sharp, which confirms that it is the Blakely decision, instead of demographic factors, that caused the discontinuity.

VI.2. Empirical Strategies and Causal Inference

VI.2.1. Empirical Models for Direct and Spillover Effects Analysis

104 I define “case” by using the most severe charge a defendant was convicted at a disposition date. See Supra Note X Yuan & Cooper (Appendix B: more details on data cleaning)
105 Age is measured by taking the difference between the date of birth and the date of the felony charge being filed.
106 Unemployment rate is the share of unemployed workers in the labor force (the sum of the employed and unemployed workers). See Bureau of Labor Statistics Website: https://www.bls.gov/cps/definitions.htm#ur
107 I calculated the residuals through a regression model which include individual and county level demographics, county fixed effects, and guidelines grid fixed effects.
The empirical analysis treats the Blakely decision as a natural experiment, which unexpectedly increased the prosecutors’ burdens of proof in seeking sentence enhancements. I employed a regression discontinuity framework to exploit the timing of the Blakely decision in order to estimate the causal effects of raising prosecutors’ burdens of proof on case outcomes.\footnote{Many economic studies adopted a similar research design to make causal inference through inter-temporal regression discontinuity frameworks. see Judson Boomhower, Drilling Like There’s No Tomorrow: Bankruptcy, Insurance and Environmental Risk, 109 American Economic Review 391, at 393 (2019) (This study examines the effects of requiring oil-drilling firms to insure against environmental risks on a wide range of outcomes related to industry composition and environmental protection. Empirically, the author adopted a regression discontinuity framework to exploit the timing of a Texas legislation which introduced such insurance requirement.) Also See Aurelie Ouss, Misaligned Incentives and the Scale of Incarceration in the United States, 191 Journal of Public Economics 104285 (2020) at 2 (Use an inter-temporal regression discontinuity framework to estimate how criminal defendants’ outcomes are affected by a 1996 California legislation that shifted the costs burden of juvenile incarceration from the state onto the county.)}

I estimate the following regression model:

\[ Y_{ijkct} = \beta_0 + \beta_1 Blakely + \beta_2 \tau + X_i \gamma + Grid_{jk} + \theta_c + \epsilon_{ijkct} \]  

(1)

\( Y_{ijkct} \) represents a set of relevant outcome variables for defendant \( i \) convicted of felony offense level \( k \) with criminal history level \( j \) in county \( c \) at time \( t \). I use defendants’ take-up of sentence enhancements to estimate the direct effects of Blakely. As to the spillover effects, we examine the following set of outcomes: (1) the level of charging offenses; (2) whether prosecutors dismissed the filed charges; (3) whether defendants plead guilty to lower-level felony; (4) whether defendants plead guilty to misdemeanors; (5) whether defendants receive incarceration sentences; and (6) defendants’ minimum sentence lengths.

The key treatment variable is \( Blakely \), an indicator of whether the case of defendant \( i \) was disposed after Blakely. It takes a value of 1 if the case is disposed of after June 2004, 0 otherwise. \( X_i \) is a vector of individual and county level demographic variables. Individual demographics include age, race, gender, criminal history points, and the types of attorney representation. County-level demographics include unemployment rate and arrest rate. \( \theta_c \) is a vector of county fixed effects, controlling for county-specific prosecution practices unobservable to researchers. I also created a vector of Guidelines grid fixed effects, \( Grid_{jk} \), by interacting convicted offense level \( j \)
and criminal history level \( k \). \( \tau_i \) is a flexible function of the running variable, month.\(^{109}\) \( \epsilon_{ijkcym} \) is an idiosyncratic error term. I cluster the standard error at month level.

The key parameter of interest is \( \beta_1 \) which captures how Blakely affects various outcomes of defendants. I expect \( \beta_1 \) to be negative if the outcome variable is defendants’ take-up of sentence enhancements. In order to interpret \( \beta_1 \) as a causal estimate, I need to impose an assumption that neither defendants nor state prosecutors manipulated the case disposition date around the timing of the Blakely decision. In other words, the timing of the Blakely decision is exogenous and uncorrelated with defendants’ underlying unobserved characteristics, \( \text{Cov}(\text{Blakely}, \epsilon_{ijkcym}) = 0 \). An important threat to identification is attorney’s potential manipulation of case disposition date.\(^{110}\) If defense attorneys anticipate the outcome of the Blakely decision, they might strategically delay the case disposition dates so that defendants can benefit from the more favorable burden of proof standard. If that is the case, we should see a jump in the number of cases after the Blakely decision. Empirically, I find no evidence suggesting that Blakely leads to meaningful changes in the number of cases decided per month.\(^{111}\) This exercise is the “McCrary test” that is widely implemented to ensure the validity of making causal inference through a regression discontinuity framework.\(^{112}\)

In addition, I employ the following event-study model to decompose the effects of Blakely into period-specific effects:

\[
Y_{ijkct} = \beta_0 + \sum_n \gamma_n \text{Blakely}_n + X_i \gamma + \text{Grid}_{jk} + \theta_c + \epsilon_{ijkc} \tag{2}
\]

Compared with Equation (1), I decomposed the Blakely into a vector of \( \text{Blakely}_n \), which means \( n \) period(s) away from the timing of Blakely decision.\(^{113}\) The time period is at the semi-year

\(^{109}\) I estimated the main regression models with different functional forms of the running variable, including month of the year fixed effects, local linear regression, and polynomial regression.
\(^{110}\) See David Lee and Thomas Lemieux, Regression Discontinuity Designs in Economics, 48 Journal of Economic Literature 281, at 283(“RD designs can be invalid if individuals can precisely manipulate the ‘assignment variable.’”)
\(^{111}\) See Supra Note X Yuan & Cooper (Figure 3.5 shows the scatter plot of monthly case volume. The Figure suggests that Blakely did not lead to meaningful changes in case volume). Also see 18-19 (more detailed discussions about why my identification is not undermine by other potential concerns, such as sample attrition or compositional changes in defendants sentenced in the same Guidelines grids.)
\(^{113}\) This type of event study analysis is first used by David Autor in assessing the effects of state courts’ enforcement of the “unjust dismissal” doctrine on the growth of the U.S. Temporary Help Service (THS) industry. See David
level. My empirical analysis includes six periods before and after the *Blakely* decision. The estimates of $\gamma_n$ recover the evolutionary path of the treatment effects and inform us of whether the effects of *Blakely* are instantaneous or lagged. I expect the effects of *Blakely* on defendants’ take-up of sentence enhancements to be instantaneous as Figure 3 strongly implies.

**VI.2.2. Empirical Models for Inter-district and Within-district Disparity Analysis**

To further examine how *Blakely* affected inter-state and within-state disparities in defendants’ take-up of sentence enhancements, I developed the following models:

\[ Y_{ijkdym} = X_i\gamma + Grid_{jk} + \delta_{ym} + \nu_{ijkdym} \]  
(3)

\[ Y_{ijkdym} = X_i\gamma + Grid_{jk} + \theta_{dy} + u_{ijkdym} \]  
(4)

Equation (3) is employed to capture inter-district disparity. The equation controls individual and county level observed characteristics, the Guidelines grid fixed effects, and year by month fixed effects. The residual term, $\nu_{ijkdym}$, captures the unexplained *inter-district* variation in the take-up of sentence enhancements among defendants who share similar characteristics and have their cases disposed in the same month. Compared with Equation (3), Equation (4) replaces the month-by-year fixed effects, $\delta_{ym}$, with district by year fixed effects, $\theta_{dy}$. The new residual term $u_{ijkdym}$ captures the unexplained *within-district* variation across different months of a year in the receipts of sentence enhancements among defendants of similar characteristics. I first perform regression analysis and numerically estimate the residual terms, $\nu_{ijkdym}$ and $u_{ijkdym}$. Then, I aggregate the estimated residuals at district-year-month level. After the data aggregation, I plot the distribution of the aggregated residual terms before and after *Blakely*. The distributional

---

Autor, Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing, 21 Journal of Labor Economics 1, at 23-24 (Constructed leads and lags of the law change to estimate how quickly the employment grew after the law change and whether the impact accelerates, stabilize, or mean reverts.)
change of the aggregated residuals reflects how Blakely affected the inter-district and within-district disparity in defendants’ take-up of sentence enhancements.

VII. Main Results for the Direct and Spillover Effects of Blakely

VII.1. The Direct Effects of Blakely on Sentence Enhancement

Table 3.2 shows the OLS results for the effects of Blakely on defendants’ take-up of sentence enhancements. I estimated five different regression specifications. The results are overall robust. The preferred specification is in Column (4) where I perform local linear regression and control demographic variables and county fixed effects. Column (5) further shows that the results are robust to including the Guidelines grid fixed effects, which compare the take-up of sentence enhancements among defendants sentenced at the same Guidelines grid pre and post Blakely.\textsuperscript{114}

In Column (4), the point estimate for the Blakely variable is around 1.8 percentage points, a 63% decrease from the pre-Blakely mean, 0.29. The estimated coefficient is statistically significant at 1% level. This result means that Blakely decreased defendants’ likelihood of receiving sentence enhancements by 63%. It also revealed that 63% of the pre-Blakely sentence enhancements were not based on strong enough evidence that can survive the beyond reasonable doubt standard. Figure 3.4 plots the event study estimates of Equation (2). The coefficients of the pre-Blakely periods are close to 0 and statistically indistinguishable from 0. The estimate decreases drastically from period 0 to period 1 and remains stable at -0.2 afterwards.

This finding can be attributed to three plausible mechanisms: changes in judicial discretion, jury factfinding, and constraint on prosecutorial discretion. I argue that it is Blakely’s constraint on prosecutorial discretion that caused the sharp drop in defendants’ take-up of sentence enhancements. First, judicial discretion is unlikely to be the mechanism because the North Carolina sentencing law does not allow judges to exercise “departure power” in making sentencing decisions.\textsuperscript{115} Judges are unable to choose sentences in aggravated ranges unless prosecutors seek

\textsuperscript{114} The spillover effects suggest that Blakely systematically changed how prosecutors process criminal cases, which inevitably lead to changes in the demographic composition of defendants sentenced at the same Guidelines grid. Comparing defendants sentenced at the same Guidelines grids pre and post Blakely are unlikely to be apple-to-apple comparison. Therefore, Column (5) is not the preferred specification.

\textsuperscript{115} Supra Note X Wright, Counting the Costs of Sentencing in North Carolina, 1980 – 2000 at 78 – 79 ("The new North Carolina system also created no departure power for judges.")
sentence enhancements and are able to establish the relevant aggravating factors. In contrast, judges at federal and many state jurisdictions exercise broad sentencing discretion in factfinding and departure. In *Blakely*, a Washington state judge enhanced the defendant’s sentence length based on his finding of “exceptional cruelty” which was never alleged by the state district attorneys. Similarly, federal judges have broad sentencing discretion and may sentence based on facts beyond what’s included in jury verdicts or plea agreements. The tight control over judicial discretion to make departure decisions in North Carolina effectively rules out the possibility that the post-*Blakely* decrease in defendants’ take-up of sentence enhancements was driven by changes in judicial discretion.

One may argue that it is the use of jury trials that explain the post-*Blakely* drop in defendants’ receipts of sentence enhancements because juries apply the stringent beyond a reasonable doubt standard to find the relevant aggravating facts. Empirical analysis does not provide support for this theory. Figure 3.5 shows the monthly share of felony cases resolved by jury trials in North Carolina. The figure shows no evidence suggesting that *Blakely* did not lead to any meaningful change in the likelihood of cases being resolved by jury trials. In addition, I estimated the *Blakely* effects using a subsample of plea bargain cases where juries play no role in determining the case outcomes. The results are the same as what is shown in Table 2.

It is the prosecutors’ behavioral response to a higher burden of proof that mostly likely caused the post-*Blakely* decrease of defendants’ receipt of sentence enhancements. This evidence supports King & Klein’s theory that enforcing the element rule gives defendants additional bargaining leverage which makes them less likely to receive sentence enhancements. Indeed, as

---

116 15A-1022.1. (a); *Id.* at 79 (North Carolina sentencing judges are able to choose a sentence length from a range. However, they are unable to go above or below the specified range in special cases.)

117 *Supra* Note X State Sentencing Guidelines, Profile and Continuum, at 19 (In Ohio, judges are free to depart from regular ranges in absence of substantial and compelling reasons and no written reasons are required for judges’ departure.);

118 *Blakely v. Washington*, 542 U.S. at 296

119 See 18 U.S.C. § 3661 ("In determining the sentence to impose within the guideline range, or whether a departure from the guidelines is warranted, the court may consider, without limitation any information concerning the background, character and conduct of the defendant, unless otherwise prohibited by law."); *United States v. Watts*, 519 U.S. 148, at 157 (Allow federal judges to sentence defendants based on conducts that jury acquitted so long as the State proves the relevant conducts by preponderance of evidence.); *See Supra* Note __ Fischman & Schanzenbach, Standard of Review and Federal Criminal Sentencing, at 411 (Sentencing judges may not accept facts stipulated in plea agreements and “proceed with sentencing on the basis of his or her own evaluation of the case.”); See 2018 United States Sentencing Commission, Guidelines Manual (2018), at 33, link: https://www.ussc.gov/sites/default/files/pdf/guidelines-manual/2018/GLMFull.pdf (Upward departure of sentence may be based on offenses that defendants are not convicted of in plea agreements.)

120 See *Supra* Note X King & Klein, at 296
my conceptual framework suggests, enforcing element rule increases the costs and decreases the expected benefits for prosecutors to pursue sentence enhancements. Prosecutors behaved as rational consumers and responded to these changes by being less likely to charge defendants with sentencing enhancements.

*Blakely* is most likely to have stopped prosecutors from pursuing sentence enhancements that require intensive evidentiary basis to establish the associated aggravating factors. The North Carolina state statute prescribes more than 20 aggravating factors to enhance defendants’ original sentence lengths. While some aggravating factors are easy to establish, others require quite intensive evidentiary basis. For example, it is straightforward to establish that the alleged offense was committed either to avoid lawful arrests or during pre-trial release periods, or the defendants were either armed or getting paid to commit the crimes. For these “objective” aggravating factors, the level of prosecutors’ burden of proof is unlikely to matter. If prosecutor can prove these objective aggravating factors by preponderance of evidence, they are also likely to prove them by beyond reasonable doubt.

On the other hand, some aggravating factors are related to defendants’ intent or motive which requires intensive evidentiary basis to establish. North Carolina state prosecutors may enhance defendants’ sentence lengths if the defendant committed hate crime, took advantage of a position of trust, or the offense is especially heinous and crucial. It is quite difficult for prosecutors to prove such motive-related aggravating factors to a jury beyond a reasonable doubt. First, prosecutors need to present sufficient circumstantial evidence to demonstrate defendants’ motives and also establish that those motives resulted in the criminal acts. Second, jurors are likely to have widely disparate standards to infer defendants’ motives based on the

---

121 A limitation of the data is that it does not include the specific information about the aggravating factors that enhanced defendants’ sentence lengths. Otherwise, I will be able to more precisely pin down which aggravating factors prosecutors decided not to prove after *Blakely*.

122 See § 15A-1340.16. (d)
123 See § 15A-1340.16. (d) (3)
124 See § 15A-1340.16. (d) (12)
125 See § 15A-1340.16. (d) (10)
126 See § 15A-1340.16. (d) (4)
127 See § 15A-1340.16. (d) (17)
128 See § 15A-1340.16. (d) (15)
129 See § 15A-1340.16. (d) (7)
130 See James Morsch, The Problem of Motive in Hate Crimes: The Argument Against Presumptions of Racial Motivation, 82 J. Crim. L. & Criminology 659, at 671 (Acknowledge that the inherent difficulty of proving motives made it difficult for prosecutors to obtain convictions under hate crime statutes.)
available evidence. Proving defendants’ motives requires a non-trial amount of evidence.\textsuperscript{131} To meet the higher “beyond a reasonable doubt” standard, prosecutors must produce substantially more evidence in proving the relevant aggravating factors. Therefore, Blakely is most likely to have deterred prosecutors from pursuing sentence enhancements based on this type of motive-related aggravating factors.

Figure 3.6 the results of heterogeneous effects of Blakely on the receipt of sentence enhancements among defendants charged with different types of charging offenses. Each bar is the estimated coefficient and the associated confidence interval of the Blakely dummy using a subsample of felony cases with the same charging offense. All the estimated coefficients are statistically significant with exceptions of forgery and homicide.\textsuperscript{132} Defendants charged with sexual assault experienced the largest decrease (4.7 percentage points) in likelihood of receiving sentence enhancements post Blakely. After accounting for the baseline probability, I found that Blakely leads to a 91% decrease in the probability of receiving sentence enhancements among defendants charged with fraud, 81% decrease among those charged with sexual assault, and 74% decrease among those charged with burglary and traffic related felony offense. On the other hand, defendants charged with homicide and robbery are only 37% and 43% less likely to receive sentence enhancements after Blakely.

There are two plausible explanations for this heterogeneity. First, it is easier to enhance defendants’ sentences when they committed severe offenses. For example, for defendants charged with homicide, there are richer facts that can enhance defendants’ sentences. Second, the asymmetry may also reflect prosecutors’ preference of punishing defendants who committed severe crimes. Due to budget constraints and limited resources, Blakely forced prosecutors to prioritize going after defendants who most deserved to be punished by sentence enhancements. The empirical analysis reflects that prosecutors’ preference of enhancing defendants’ sentence lengths when they committed relatively more severe offenses, such as homicide and robbery.

I further examine the heterogenous effects of Blakely on defendants’ likelihood of receiving sentence enhancements by the types of attorney representations. Gideon v. Wainwright

\textsuperscript{131} Id. at 665 (In civil suits, victims experienced difficulty in proving defendants’ discriminatory motives by a preponderance of evidence standard and were unable to recover damages.)

\textsuperscript{132} For the subsample of cases with homicide charge, the estimated coefficient for the Blakely dummy is -0.46 and the standard error is 0.029. I did not include this coefficient in Figure 3.6 because the wide confidence interval makes it harder to read other coefficients.
held that the Sixth Amendment requires all states to provide defendants with legal counsels in all
criminal cases.\textsuperscript{133} There is a growing empirical literature rigorously assessing the effectiveness of
legal representation on case outcomes of criminal defendants. Yotam Shem-tov estimated the
causal effects of public defender representation on defendants’ case outcomes by exploiting the as
good as random assignment of public defenders among criminal cases of multiple defendants.
Shem-tov found that, compared with being represented by Court’s appointed counsels, defendants
represented by public defenders are 6\% less likely to be convicted and 22\% less likely to receive
prison sentences.\textsuperscript{134} Many other similar studies also show that public defender representation leads
to more favorable outcomes to criminal defendants than private attorney.

Legal representation is a critical check on prosecutors’ arbitrary decisions to enhance
defendants’ sentences. Effective criminal defense attorneys will more carefully scrutinize
prosecutors’ cases for sentence enhancements, point out the evidentiary flaws, and help defendants
avoid longer sentences. In particular, \textit{Blakely} provides a unique opportunity to examine which type
of attorney is more effective to have defendants benefit from a higher burden of proof standard
imposed on the state prosecutors by the new law.

My data includes the following four types of attorney representation: private attorney,
public defender, Court appointed counsels, or \textit{pro se} defense. Under the North Carolina state law,
divident defendants are entitled to legal representation by either public defendants or Court’s
appointed counsels.\textsuperscript{135} Private attorney only represents non-indigent defendants. Both indigent and
non-indigent defendants may waive their right to counsel upon proper legal procedures.\textsuperscript{136} In my
analysis sample, 49\% of the defendants are represented by court-appointed counsels, 22\% by
public defenders, and 24\% by private attorneys. The rest 5\% of the defendants chose to waive their
right to counsel. I estimated the effects \textit{Blakely} on defendants’ likelihood of receiving sentence
enhancements using different subsamples based on the type of defendants’ attorney representation.

Figure 3.7 shows the results. Each bar represents the estimated coefficient and the
associated 95\% confidence interval for the \textit{Blakely} variable using different subsamples based on
the type of defendants’ attorney representation. For indigent defendants represented by public

---

\textsuperscript{133} 372 U.S. 335 (1963)
\textsuperscript{134} See Yotam Shem-tov, Make-or-Buy? The Provision of Indigent Defense Service in the U.S., Review of
Economics and Statistics 1 (2020), at 1
\textsuperscript{135} § 7A-450 (b)
\textsuperscript{136} § 7A-457. Waiver of counsel; pleas of guilty. § 15A-603. Assuring defendant's right to counsel
defenders, *Blakely* decreased their likelihood of receiving sentence enhancements by 2.1 percentage points, a 62% decrease from the pre-*Blakely* mean 0.034. In contrast, the decrease is much smaller for other indigent defendants represented by Court appointed counsels. The blue bar shows that the point estimate is 1.4 percentage points, a 47% decrease from the pre-*Blakely* mean, 0.3. This contrast suggests that court appointed counsels are much less effective to help defendants avoid sentence enhancements after *Blakely*. This finding is consistent with the causal evidence in the current empirical literature that public defenders deliver favorable outcomes to criminal defendants than Court appointed counsels.137

Additionally, for the defendants represented by private attorneys, their likelihood of receiving sentence enhancements decreased by 1.7 percentage points, a 65% decrease from the pre-*Blakely* mean, 0.26. This result suggests that private attorneys are as effective as public defenders in helping defendants avoid sentence enhancements. Lastly, the yellow bar suggests that *Blakely* did not lead to meaningful changes in the likelihood of receiving sentence enhancements among defendants who waived their right to counsel. The estimated coefficient has small magnitude, less than negative half percentage point, and is statistically indistinguishable from 0. This set of analysis indicates that imposing a higher burden of proof on prosecutors does not automatically lead to fewer sentence enhancements. Defense attorneys’ rigorous advocacy is an important factor that explains the post-*Blakely* drop in defendants’ take-up of sentence enhancements shown in Figure 3.3. Defense attorneys may have forced prosecutors to seriously reconsider whether the evidentiary basis for sentence enhancements can survive a proof by a beyond reasonable doubt, which changed prosecutors’ decisions to enhance defendants’ sentences.

**VII.2. The Spillover Effects of *Blakely* on Plea Bargain Outcomes**

This set of analysis investigate whether the additional procedural barrier to pursue sentence enhancements lead to systematic changes in how prosecutors choose to process criminal cases, which I defined as the “spillover effects” of *Blakely*. Such spillover effects are plausible considering the prosecutors’ broad discretion in making a wide range of prosecution related decisions. Additionally, the theoretical analysis from Stephanos Bibas indicates that prosecutors are likely to treat defendants more harshly in response to the Courts’ enforcement of the element

---

137 See *Supra* Note X
rule. My empirical analysis also directly tests the validity of Bibas’s theoretical predictions on prosecutors’ behavioral changes in response to Court’s enforcement of the element rule.

I measure the spillover effects by examining the following case outcomes: (1) the number of charges prosecutors brought against defendants; (2) whether prosecutors dismiss filed charges; (3) whether a case is resolved by plea bargain; (4) whether a defendant pleaded guilty to a lower felony offense; (5) whether a defendant pleaded guilty to a misdemeanor offense; and (6) the length of minimum sentences. Table 3.3 reports the results of the spillover effects.

In Column (1), the outcome variable is the total number of charges associated with each case. I limit the sample to cases with less than 10 charges so that the results are less likely to be driven by outliers. The estimated coefficient is 0.12 and statistically significant at 1% level. The point estimate implies that after Blakely prosecutors increased the number of charges by 4.8%. An earlier empirical study finds that higher number of charging counts is correlated with longer sentences. This finding suggests that sentence-maximizing prosecutors increased the number of charges to compensate the additional procedural barrier of seeking sentence enhancements after Blakely.

In Column (2), the outcome variable whether prosecutors dismissed a felony case. In this specification, I did not control individual and county level characteristics as I do not observe that information for dismissed cases. The regression model only includes a Blakely dummy, county fixed effects, and month of year fixed effects. The point estimate suggests that Blakely increased the likelihood of case dismissal by 1.46 percentage points, a 15.8% increase from the pre-intervention mean. The estimated coefficient is statistically significant at 1% level. The effect size is quite large. This finding implies that the high costs of complying with the jury factfinding requirement significantly changed prosecutors’ discretionary decisions on whether to prosecute given criminal cases. Blakely might have forced prosecutors to dismiss marginally prosecutable cases and to concentrate limited resources on prosecuting defendants eligible for sentence enhancements.

Then, I investigate whether Blakely have changed the likelihood of cases resolved by plea bargain. The result is shown is Column (3) of Table 3.3. The estimated coefficient has small magnitude, less than half percentage point. Despite the statistical significance, the estimated

---

138 Keith A. Wilmot & Cassia Spohn, Real-Offense Sentencing: An Analysis of Relevant Conduct Under the Federal Sentencing Guidelines, 15 Criminal Justice Policy Review 324, 332-333 (Defendants charged with more than one count received 6 months longer sentences than those charged with only one count.)
change is not economically meaningful as the pre-Blakely plea bargain rate is 95.7%. The result implies that Blakely did not significantly change the plea bargain rate. Legal scholars argued that plea bargain does not occur in the shadow of substantive criminal law. Instead, plea bargaining is primarily driven by non-legal forces, such as prosecutors’ preference and budget constraints. Blakely represents a significant modification of substantive sentencing law, which made it more difficult for prosecutors to win sentence enhancements at trials. As a result, one might anticipate sizable changes in plea bargain rate. However, the empirical finding shows that plea bargain is almost unaffected by the change in substantive criminal law.

Furthermore, I investigate the effects of Blakely on charge movements which refer to the difference between the defendants’ filing and convicted charges. An empirical study by Ronald Wright and Rodney Engen analyzed the data of North Carolina criminal prosecution and documented the prevalence of charge reduction, from felony to either less severe felony or misdemeanors. Prosecutors’ exercise of discretion is likely to be the main mechanism underlying the frequent charge movements, particularly in jurisdictions with mandatory sentencing guidelines like North Carolina. Therefore, charge movement is another important outcome largely driven by prosecutors’ exercise of discretion.

The results are in Column (4) and (5). The outcome variable in Column (4) specification is whether a defendant pleaded guilty to lesser felony offense and the estimated coefficient is -.011, statistically significant at 1% level. The estimate suggests that Blakely decreased defendants’ likelihood of pleading guilty to lesser felony offense by 1.1 percentage points, an 8.3% decrease from the pre-Blakely mean. Column (5) shows the effects of Blakely on the likelihood of defendants pleading guilty from felony to misdemeanor. The point estimate is 0.009 and

---

139 See Stephanos Bibas, Plea Bargaining Outside the Shadow of Trial, 117 Harvard Law Review 2463 (2004), 2467 (“many plea bargains diverge from the shadows of trials.”); See William Stunz, Plea Bargaining and the Criminal Law’s Disappearing Shadow, 117 Harvard Law Review 2548, 2548 (“the law’s effect on plea bargaining is much smaller than conventional wisdom would have it.”)
140 Id, Stunz, at 2550 (for some crime, “plea bargains take place in the shadow of prosecutors’ preference, voters’ preference, budget constraints, and other forces – but not in the shadow of the law.”)
141 Ronald F. Wright & Rodney L. Engen, The Effects of Depth and Distance in a Criminal Code on Charging, Sentencing, and Prosecutor Power, 84 N.C. L. Rev. 1935 (2006), 1957 – 1958 (Document extensive charge reductions across all different classes of charged offense. The reductions are higher among the most severe offenses where prison sentences are mandatory.)
142 Supra Note X Barkow, Recharging the Jury, at 98-100 (Argued that prosecutors can more effectively pressure defendants to plead guilty under the guidelines sentencing regime where the risk of receiving longer sentences at trials is much higher than discretionary sentencing regime.)
statistically significant at 1% level, suggesting that Blakely increased defendants’ probability of pleading guilty to misdemeanors by 0.9 percentage point, a 5.4% increase.

The findings shed light on how state prosecutors under the constraints of limited resources modified the plea bargain practice in response to Blakely. First, prosecutors compensate the additional procedural barrier of pursuing sentence enhancements by being less willing to lower some initial charges to less severe felony offenses. However, keeping the conviction the same as the charging offense is not cost free. Prosecutors are likely to face pressure from the defense side to produce sufficient evidence to justify their decisions. As a result, prosecutors became more willing to lower some felony offense to misdemeanors. In this way, prosecutors can allocate resources to convict some defendants as charged.

Column (6) shows the effects of Blakely on defendants’ likelihood of pleading to incarceration punishment. In developing the empirical analysis, I impose two restrictions on the analysis sample. First, I restricted the convicted offenses to Classes E, F, G, H & I. The sentencing guidelines permit non-incarceration sentences for individuals convicted of this group of offense. Second, I focus on plea bargain cases. The estimated coefficient is 0.013 and statistically significant at 1% level, suggesting that Blakely increased defendants’ likelihood of pleading to incarceration punishment by 1.3 percentage points, an 4.3% increase.

Under the North Carolina law, the Court makes final decisions on whether defendants are punishable by imprisonment. In plea bargained cases, prosecutors’ recommendations may significantly influence the courts’ decisions. The post-Blakely changes in incarceration rate is attributable to the changing behavior from both judges and prosecutors. Blakely transferred the authority of finding aggravating factors from judges to juries in North Carolina. It is plausible that judges may react to the loss of factfinding power by becoming harsher to defendants who committed crimes in aggravated manners, which leads to a higher incarceration rate. Alternatively, prosecutors may have responded to Blakely by pursuing incarceration punishment more aggressively to circumvent the procedural barrier of punishing defendants by sentence.

---

143 According to the sentencing guidelines in Figure 3.2, defendants will receive an incarceration sentence, or “active punishment,” if they are convicted of an offense of class E, or higher.
144 Supra Note X, at 27 (The state courts have the discretion to choose appropriate punishment if the guidelines prescribe multiple dispositions for the convicted defendants.)
enhancements. Therefore, prosecutors and judges are likely to have jointly contributed to the sizable increase in defendants’ incarceration rate after Blakely.  

Finally, Column (6) shows the results for the effects of Blakely on defendants’ minimum sentence length. The result implies that Blakely did not significantly change defendants’ minimum sentence lengths as the point estimate has a small magnitude and is statistically indistinguishable from 0. Earlier results show that Blakely caused a large decrease, about 63%, in defendants’ take-up of sentence enhancements, which should have mechanically decreased defendants’ minimum sentence lengths. Therefore, the finding on minimum sentence length is contrary to my expectation.

There are two plausible explanations for this surprising finding. One is the compositional changes of prosecuted cases. After Blakely, prosecutors focus on prosecuting more severe cases that carry longer sentences, which offset the anticipated sentence reduction associated with jury factfinding rule. This explanation is consistent with my earlier finding that Blakely decrease the likelihood of case dismissal by about 16%. Also, the null finding suggests that judges and prosecutors may have jointly circumvented the new procedural protection of defendants against longer sentences. According to the North Carolina state law, judges choose defendants’ appropriate minimum sentence lengths from the applicable ranges according to the sentencing guidelines. If judges cannot easily move up the range, they might choose increase defendants’ regular sentence lengths within the applicable ranges, which is within their discretion. Defendants’ sentence lengths also reflect various prosecutors’ decisions including charging, prosecution, and plea bargain. The abundant evidence on prosecutors “compensating behavior” in making prosecution-related decisions can also explain the null finding of the minimum sentence lengths.

VIII. Results on Inter-District and Within-District Disparities

So far, the empirical analysis primarily focuses on mean changes in various defendants’ outcomes. This set of analysis explore how Blakely changed the variance of defendants’ take-up of sentence enhancements. The low preponderance of evidence standard tends to result in a larger

\[ \text{Due to data limitation, I am unable to estimate how prosecutors and judges separately contribute to the increase in defendants’ incarceration rate.} \]

\[ \text{I do not observe defendants’ final sentence length, which would be a valuable outcome to analyze.} \]
variance in defendants’ take-up of sentence enhancements for two reasons. First, prosecutors have
the complete discretion to enhance defendants’ sentences. There is little institutional oversight to
prevent prosecutors’ abuse of discretion. Second, since preponderance of evidence is an easy
standard to meet, prosecutors may pursue sentence enhancements based on their tastes rather than
evidentiary strength. Prosecutors’ taste is a function of various factors, such as prosecutors’
experience, professional goals, and discriminatory intent. Defendants of similar characteristics
may receive sentence enhancements with different probabilities due to assigned prosecutors.

Blakely shifted prosecutors’ burden of proof to a beyond reasonable doubt standard. Meeting
the higher burden of proof standard usually requires more intensive evidentiary basis, which might
be hard to establish. Blakely forced prosecutors to prioritize considering the evidentiary strength
in determining whether to enhance defendants’ sentences, which substantially constrained
prosecutors’ opportunities to act arbitrarily or to maximize idiosyncratic preferences. As a result,
I anticipate that Blakely will reduce the variance of defendants’ take-up of sentence enhancements.
A simple descriptive empirical analysis confirms my expectation of post-Blakely changes in the
variance of defendants’ take-up of sentence enhancements. Figure 3.8 shows the distributions of
monthly sentence enhancements rates before and after Blakely. Blakely not only shifted the
distribution to the left but also changed the second moment. The pre-Blakely distribution is more
spread out. In contrast, the post-Blakely distribution has noticeably smaller variance and is more
centered around the mean.

Then I estimated changes in the inter-district and inter-district disparities of sentence
enhancements rates. I measure the disparities by examining the variation of the estimated residuals
of Equation (3) and (4). I adopted two separate approaches to demonstrate how Blakely changed
the disparities of sentence enhancements. First approach is to separately estimate the density of the
residualized sentence enhancement rates in pre and post Blakely periods. Alternatively, I also use
boxplots to show the distributional change in the residualized sentence enhancement rates.

Figure 3.9 shows the results of how Blakely changed the inter-district disparity of defendants’
sentence enhancement rate. The estimated residuals from Equation (3) capture the unexplained
inter-district variation of sentence enhancements rates within a given month among defendants

---

147 In this empirical exercise, I first aggregate the individual level sentence enhancement dummy at the monthly
level. Then, I separately plotted the distribution of the aggregated monthly sentence enhancement rates in pre and
post Blakely periods.
with similar observable characteristics. Figure 3.9(A) shows that the distribution of the residuals becomes much less spread-out after Blakely, suggesting that constraining prosecutorial discretion reduced inter-district disparity of sentence enhancements. Boxplots in Figure 3.9(B) reaffirms this finding. The upper and lower end of the box capture the seventy-fifth and twenty-fifth percentile mean district effects, also known as the inter-quartile range (IQR). Defendants prosecuted at the districts with higher percentile mean district effects are more likely to receive sentence enhancements. The boxplots suggest that the IQR became substantially narrower in the post Blakely periods. It is worth noting that Blakely did not eliminate the inter-district disparity of sentence enhancement rates. There are several factors that can plausibly explain the remaining inter-district disparity, such as the differences in defendants’ demographics across districts or the differences in the prosecution priorities at different prosecutors’ offices.

Figure 3.10 shows the results of within-district disparity changes. The distribution of residuals from Equation (4) captures the within-district disparity in sentence enhancement, referring the unexplained variation within a district across different months of a given year. The residuals do not reflect inter-district variation because the regression model include district by year fixed effects. Figure 3.10(A) shows the distributional changes in within-district disparity pre and post Blakely. The residual distribution during the post-Blakely periods is much less spread-out and more centered around the mean. Figure 3.10(B) shows the boxplots. The box represents the IQR of the mean month effects. Defendants are more likely to receive sentence enhancements if they are prosecuted in a month with higher percentile mean month effects. The graph shows that the IQR is much narrower after Blakely, suggesting that constraining prosecutorial discretion has clear negative effects in reducing within district disparity.

IX. Conclusion

This article examines the direct and spillover effects of constraining prosecutor’s discretion to enhance defendants’ sentence lengths. My analysis uses detailed criminal case records from the North Carolina Administrative Office of Courts. I employed a regression discontinuity empirical

---

148 A better measure of within-district disparity is to examine the inter-prosecutor disparity within a district, meaning the disparity in sentence enhancement rate across prosecutors working in the same district. Unfortunately, I am unable to implement this empirical exercise because my data does not contain information related the assigned prosecutors in each criminal case.
framework to exploit the timing a US Supreme Court decision, *Blakely v. Washington*, which more rigorously enforced Sixth Amendment Right to Jury and increased the North Carolina state prosecutors’ burden of proof in successfully seeking enhancements from preponderance of evidence to beyond a reasonable doubt. My analysis offers new empirical insights partially clarifying a theoretical debate among legal scholars on whether enforcing the element rule makes defendants better off.

I first showed that *Blakely* had the direct effects in reducing defendants’ likelihood of receiving sentence enhancements by 63%. The large decrease is primarily driven by prosecutors’ decisions of not pursuing sentence enhancements. The heterogeneous analysis of the direct effects show that the decrease is larger among defendants represented by public defenders and those convicted with less severe crimes. The direct effects results are consistent with the theory from King & Klein that the expensive jury trial procedure strengthens defendants’ bargaining position, which helped them avoid sentence enhancements.

I then examined how *Blakely* affects other prosecution related decisions, the spillover effects. Consistent with Bibas’s theory, I find evidence suggesting that prosecutors treated defendants more harshly in response to the additional procedural barrier of punishing defendants with sentence enhancements. Specifically, *Blakely* resulted in a 4.8% increase in the number of charges filed against defendants, 8.3% decrease in the probability of pleading guilty to lesser felony, and 4.3% increase in the likelihood of receiving imprisonment punish in plea bargains. In addition, I also find evidence suggesting that the expensive jury process forced prosecutors focus on prosecuting more severe criminal offenses by dismissing more felony cases and being more willing to lower convicted offense from felony to misdemeanor. After *Blakely*, there was a 15.8% increase in cases that prosecutors declined to prosecute and 5.4% increase in the likelihood of pleading guilty to misdemeanors. Overall, I do not find evidence suggesting that *Blakely* significantly changed defendants’ minimum sentence lengths. This result implies that the new procedural protection of defendants against longer sentence is circumvented by prosecutors’ exercise of discretion in making a wide range of prosecution decisions.

Finally, I also showed the constraining prosecutors’ discretion also changed the variance of sentence enhancements. *Blakely* decreased the unexplained inter-district and within-district disparities by almost 50%. This result highlight that unwarranted disparities are artifacts of excessive discretion which enable prosecutors to make arbitrary decisions.
Figures and Tables

Figure 3.2: Mandatory Sentencing Guidelines in North Carolina

<table>
<thead>
<tr>
<th>OFFENSE CLASS</th>
<th>I 0 Points</th>
<th>II 1-4 Points</th>
<th>III 5-9 Points</th>
<th>IV 9-14 Points</th>
<th>V 15-19 Points</th>
<th>VI 19+ Points</th>
<th>DEPOSITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>Disposition</td>
</tr>
<tr>
<td>B1</td>
<td>260-300</td>
<td>285-350</td>
<td>336-420</td>
<td>384-480</td>
<td>384-480</td>
<td>384-480</td>
<td>Presumptive Range</td>
</tr>
<tr>
<td>C</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>Disposition</td>
</tr>
<tr>
<td>E</td>
<td>55-73</td>
<td>90-100</td>
<td>93-116</td>
<td>107-133</td>
<td>121-151</td>
<td>135-168</td>
<td>Mitigating Range</td>
</tr>
<tr>
<td>F</td>
<td>64-28</td>
<td>60-30</td>
<td>70-93</td>
<td>80-107</td>
<td>90-121</td>
<td>101-135</td>
<td>Disposition</td>
</tr>
<tr>
<td>G</td>
<td>55-80</td>
<td>71-95</td>
<td>107-179</td>
<td>135-146</td>
<td>159-167</td>
<td>166-181</td>
<td>Presumptive Range</td>
</tr>
<tr>
<td>H</td>
<td>81-86</td>
<td>61-77</td>
<td>82-103</td>
<td>94-117</td>
<td>107-133</td>
<td>117-146</td>
<td>Mitigating Range</td>
</tr>
<tr>
<td>I</td>
<td>88-111</td>
<td>60-40</td>
<td>70-54</td>
<td>80-107</td>
<td>90-121</td>
<td>101-135</td>
<td>Disposition</td>
</tr>
</tbody>
</table>

A: Actual Punishment  I: Informative Punishment  C: Community Punishment
Punishments shown are in months and represent the range of minimum sentences.

Effective for Offenses Committed on or after 12/1/95 Through 11/30/09

FELONY PUNISHMENT CHART

PRIORITY RECORD LEVEL

DISPOSITION

Presumptive Range

Mitigating Range

72
Figure 3.3: Monthly Average Rate of Cases with Sentence Enhancements

A: Raw Data

B: Residuals

Note: Each dot represents the monthly share of cases where defendants receive sentence enhancement. Then I overlay the scatterplot with curves fitted using the second-degree polynomials.
Figure 3.4: Results of Event Study Estimates

Note: The graph plots the event study estimates of Equation (2) with the 95% confidence intervals. The regression model includes a rich set of individual and county level demographic controls, county fixed effects, and the Guidelines grid fixed effects. The standard error is clustered at month level.
Figure 3.5: Monthly Average of Cases Resolved by Jury Trials

Note: The sample period is from 2001 to 2007. Each dot represents a monthly share of cases resolved by jury trials. The shares are calculated using the raw data.
Figure 3.6: The Heterogenous Effects of *Blakely* on Sentence Enhancements (by Offense Type)

Note: The sample period is from 2001 to 2007. Each bar represents the estimated coefficient and the associated confidence interval from a *local linear regression* model using different subsamples based on the type of charging offense. The regression model includes individual and county level control variables and county fixed effects. The standard error is clustered at monthly level.
Figure 3.7: The Effects of *Blakely* on Sentence Enhancements (By Representation Types)

Note: The sample period is from 2001 to 2007. Each bar represents the estimated coefficient and the associated confidence interval from a *local linear regression* model using different subsamples based on the type of defendants’ attorney representation. The regression model includes individual and county level control variables and county fixed effects. The standard error is clustered at monthly level.
Figure 3.8 Distribution of Monthly Sentence Enhancement Rate Pre and Post Blakely
Figure 3.9: The Effects of *Blakely* on Inter-District Disparity

Panel (A)

Panel (B)

Note: The sample period is from 2001 to 2007.
Figure 3.10: The Effects of *Blakely* on Within-District Disparities

Panel (A)

Panel (B)

Note: The sample period is from 2001 to 2007.
Table 3.1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.84</td>
<td>0.37</td>
</tr>
<tr>
<td>White</td>
<td>0.40</td>
<td>0.49</td>
</tr>
<tr>
<td>Black/Hispanic</td>
<td>0.57</td>
<td>0.49</td>
</tr>
<tr>
<td>Age</td>
<td>30.10</td>
<td>10.35</td>
</tr>
<tr>
<td>Represented by Private Attorney</td>
<td>0.24</td>
<td>0.43</td>
</tr>
<tr>
<td>Represented by Public Defender</td>
<td>0.22</td>
<td>0.41</td>
</tr>
<tr>
<td>Criminal History Points</td>
<td>4.60</td>
<td>5.30</td>
</tr>
<tr>
<td>Arrest Rate</td>
<td>535.98</td>
<td>221.07</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>5.81</td>
<td>1.57</td>
</tr>
<tr>
<td>Plea Bargain</td>
<td>0.86</td>
<td>0.34</td>
</tr>
<tr>
<td>Plead Guilty to Lower Felony</td>
<td>0.11</td>
<td>0.31</td>
</tr>
<tr>
<td>Plead Guilty to Misdemeanors</td>
<td>0.15</td>
<td>0.35</td>
</tr>
<tr>
<td>Aggravated</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>Mitigated</td>
<td>0.08</td>
<td>0.27</td>
</tr>
<tr>
<td>Minimum Sentence Length</td>
<td>17.37</td>
<td>39.44</td>
</tr>
<tr>
<td>Incarceration</td>
<td>0.31</td>
<td>0.46</td>
</tr>
</tbody>
</table>

The Number of Observations 213261

Note: The analysis sample combines data from the following three sources: North Carolina Administrative Office of Courts (AOC), Bureau of Labor Statistics (BLS), and Uniform Crime Report (UCR). Arrest rate is the number of arrests made per 100,000 county population. The sample period is from 2001 to 2007.
Table 3.2: The Effects of Blakely on Defendants’ Take-up of Sentence Enhancements

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blakely</td>
<td>-0.0153***</td>
<td>-0.0157***</td>
<td>-0.0181***</td>
<td>-0.0182***</td>
<td>-0.0180***</td>
</tr>
<tr>
<td></td>
<td>(0.000789)</td>
<td>(0.00112)</td>
<td>(0.00126)</td>
<td>(0.00124)</td>
<td>(0.00124)</td>
</tr>
<tr>
<td>Local Linear Regression</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Demographic Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Guidelines Grid Fixed Effects</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>213261</td>
<td>213261</td>
<td>186813</td>
<td>186813</td>
<td>186813</td>
</tr>
</tbody>
</table>

Note: the Pre-Blakely mean of sentence enhancements is 0.029. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The sample period is from 2001 to 2007. The analysis sample includes all the criminal cases where defendants are charged from felony from 2001 to 2007 in North Carolina. The standard error is clustered at month level. Significance level: * p<0.1, ** p<0.05, *** p<0.01.
Table 3.3: The Spillover Effects of Blakely on Defendants’ Case Outcomes

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Charges</td>
<td>Blakely</td>
<td>Dismissed by Prosecutors</td>
<td>Plea Bargain</td>
<td>Plead Guilty to Lower Felony</td>
<td>Plead Guilty to Misdemeanors</td>
<td>Incarceration</td>
<td>Minimum Sentence Length</td>
</tr>
<tr>
<td></td>
<td>0.119***</td>
<td>0.0146***</td>
<td>0.00363***</td>
<td>-0.0114***</td>
<td>0.00895***</td>
<td>0.0131***</td>
<td>0.134</td>
</tr>
<tr>
<td></td>
<td>(0.0169)</td>
<td>(0.00159)</td>
<td>(0.00106)</td>
<td>(0.00273)</td>
<td>(0.00280)</td>
<td>(0.00298)</td>
<td>(0.159)</td>
</tr>
<tr>
<td>Pre-Blakely Means</td>
<td>2.475</td>
<td>0.092</td>
<td>0.957</td>
<td>0.133</td>
<td>0.168</td>
<td>0.299</td>
<td>23.195</td>
</tr>
<tr>
<td>R Squared</td>
<td>0.090</td>
<td>0.053</td>
<td>0.033</td>
<td>0.127</td>
<td>0.134</td>
<td>0.279</td>
<td>0.209</td>
</tr>
<tr>
<td>N</td>
<td>181060</td>
<td>213261</td>
<td>186813</td>
<td>182779</td>
<td>182779</td>
<td>135414</td>
<td>184791</td>
</tr>
</tbody>
</table>

Note: The data comes from the North Carolina Administrative Office of Courts (AOC). The sample period is from 2001 to 2007. With an exception to Column (2), the estimate in each column comes from a regression model which regresses the specified outcome variable on a Blakely dummy, a rich set observed control variables, county fixed effects, and month of year fixed effects. The standard errors are clustered at month level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Appendix

1.A Additional Results on the Effects of Each Type of the Laws

This Appendix reports the more detailed results on how cash and medical benefit payments are affected by each of the five different types of law we empirically examine. For each type of the laws, we show the results of its effects on cash and medical benefit payments across different regression models with state fixed effects, year fixed effects and state-specific time trend. Overall, the results are fairly robust across various specifications.

1.A.1 Laws Prohibiting Liberal Construction Rule

Table 1.A.1: The Effects of Liberal Construction Bans on Cash Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Benefits Index (log)</td>
<td>0.141</td>
<td>-1.461</td>
<td>1.845</td>
<td>1.725</td>
<td>1.167</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.398)</td>
<td>(0.423)</td>
<td>(0.402)</td>
<td>(0.212)</td>
</tr>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>-0.124</td>
<td>0.0677*</td>
<td>-0.0134</td>
<td>-0.0482</td>
<td>-0.00924</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
<td>(0.0381)</td>
<td>(0.0467)</td>
<td>(0.0471)</td>
<td>(0.0212)</td>
</tr>
<tr>
<td>Liberal Construction Ban</td>
<td>0.0468</td>
<td>-0.320**</td>
<td>-0.233*</td>
<td>-0.227*</td>
<td>-0.0931*</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td>(0.153)</td>
<td>(0.116)</td>
<td>(0.117)</td>
<td>(0.0538)</td>
</tr>
<tr>
<td>State Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>States without the Laws Before 1997</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State-specific Time Trend</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.020</td>
<td>0.232</td>
<td>0.396</td>
<td>0.418</td>
<td>0.747</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>860</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of cash benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 1.A.2: The Effects of Liberal Construction Ban on Medical Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>0.215**</td>
<td>0.132***</td>
<td>-0.00869</td>
<td>-0.0283</td>
<td>0.0181</td>
</tr>
<tr>
<td></td>
<td>(0.0803)</td>
<td>(0.0439)</td>
<td>(0.0414)</td>
<td>(0.0437)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>Liberal Construction Ban</td>
<td>0.100</td>
<td>-0.197***</td>
<td>-0.0850**</td>
<td>-0.0709</td>
<td>-0.0404</td>
</tr>
<tr>
<td></td>
<td>(0.0825)</td>
<td>(0.0340)</td>
<td>(0.0423)</td>
<td>(0.0443)</td>
<td>(0.0274)</td>
</tr>
<tr>
<td>State Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>States without the Laws Before 1997</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State-specific Time Trend</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.147</td>
<td>0.073</td>
<td>0.306</td>
<td>0.343</td>
<td>0.638</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>860</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of medical benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
1.A.2 Laws Presuming Intoxication as the Proximate Cause of Injuries

Table 1.A.3: The Effects of Intoxication Laws on Cash Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Benefits Index (log)</td>
<td>0.113</td>
<td>-1.540***</td>
<td>1.856***</td>
<td>1.882***</td>
<td>1.203***</td>
</tr>
<tr>
<td></td>
<td>(0.424)</td>
<td>(0.376)</td>
<td>(0.522)</td>
<td>(0.623)</td>
<td>(0.219)</td>
</tr>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>-0.0888</td>
<td>0.0671*</td>
<td>-0.0175</td>
<td>-0.0343</td>
<td>-0.00799</td>
</tr>
<tr>
<td></td>
<td>(0.0944)</td>
<td>(0.0383)</td>
<td>(0.0484)</td>
<td>(0.0488)</td>
<td>(0.0215)</td>
</tr>
<tr>
<td>Intoxication Laws</td>
<td>-0.257*</td>
<td>-0.241**</td>
<td>-0.185*</td>
<td>-0.192*</td>
<td>-0.157*</td>
</tr>
<tr>
<td></td>
<td>(0.129)</td>
<td>(0.103)</td>
<td>(0.101)</td>
<td>(0.105)</td>
<td>(0.0790)</td>
</tr>
<tr>
<td>State Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>States without the Laws Before 1997</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>State-specific Time Trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.059</td>
<td>0.214</td>
<td>0.390</td>
<td>0.391</td>
<td>0.751</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>860</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of cash benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>0.253***</td>
<td>0.130***</td>
<td>-0.00844</td>
<td>-0.0261</td>
<td>0.0193</td>
</tr>
<tr>
<td></td>
<td>(0.0754)</td>
<td>(0.0428)</td>
<td>(0.0405)</td>
<td>(0.0402)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>Intoxication Laws</td>
<td>-0.177*</td>
<td>-0.307**</td>
<td>-0.250**</td>
<td>-0.254**</td>
<td>-0.129*</td>
</tr>
<tr>
<td></td>
<td>(0.0949)</td>
<td>(0.135)</td>
<td>(0.119)</td>
<td>(0.119)</td>
<td>(0.0695)</td>
</tr>
<tr>
<td>State Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>States without the Laws Before 1997</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State-specific Time Trend</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.172</td>
<td>0.117</td>
<td>0.346</td>
<td>0.350</td>
<td>0.642</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>860</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of medical benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
### 1.A.3 Benefits Apportionment Laws Due to Preexisting Conditions

Table 1.A.5: The Effects of Apportionment Laws on Cash Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Benefits Index (log)</td>
<td>0.136</td>
<td>-1.556***</td>
<td>1.875***</td>
<td>1.612***</td>
<td>1.203***</td>
</tr>
<tr>
<td></td>
<td>(0.455)</td>
<td>(0.406)</td>
<td>(0.500)</td>
<td>(0.330)</td>
<td>(0.209)</td>
</tr>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>-0.113</td>
<td>0.0617</td>
<td>-0.0218</td>
<td>-0.0235</td>
<td>-0.00952</td>
</tr>
<tr>
<td></td>
<td>(0.0982)</td>
<td>(0.0373)</td>
<td>(0.0483)</td>
<td>(0.0612)</td>
<td>(0.0213)</td>
</tr>
<tr>
<td>Apportionment Laws</td>
<td>-0.0402</td>
<td>-0.265***</td>
<td>-0.195***</td>
<td>-0.212***</td>
<td>-0.266**</td>
</tr>
<tr>
<td></td>
<td>(0.127)</td>
<td>(0.0565)</td>
<td>(0.0549)</td>
<td>(0.0607)</td>
<td>(0.112)</td>
</tr>
</tbody>
</table>

|                          | X       | X       | X       | X       | X       |
| State Fixed Effects      |         |         |         |         |         |
| Year Fixed Effects       | X       | X       | X       |         |         |
| States without the Laws Before 1997 | X       |         |         |         |         |
| State-specific Time Trend |         |         |         |         |         |
| R-Squared                | 0.021   | 0.204   | 0.382   | 0.390   | 0.751   |
| N                        | 1000    | 1000    | 1000    | 580     | 1000    |

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of cash benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 1.A.6: The Effects of Apportionment Laws on Medical Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>0.232***</td>
<td>0.133***</td>
<td>-0.0120</td>
<td>-0.0453</td>
<td>0.0181</td>
</tr>
<tr>
<td></td>
<td>(0.0801)</td>
<td>(0.0444)</td>
<td>(0.0415)</td>
<td>(0.0508)</td>
<td>(0.0216)</td>
</tr>
<tr>
<td>Apportionment Laws</td>
<td>0.0311</td>
<td>-0.165**</td>
<td>-0.0962*</td>
<td>-0.101</td>
<td>-0.266***</td>
</tr>
<tr>
<td></td>
<td>(0.0763)</td>
<td>(0.0642)</td>
<td>(0.0563)</td>
<td>(0.0631)</td>
<td>(0.0309)</td>
</tr>
<tr>
<td>State Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>States without the Laws Before 1997</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>State-specific Time Trend</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.138</td>
<td>0.058</td>
<td>0.305</td>
<td>0.274</td>
<td>0.645</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>580</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of medical benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
### 1.A.4 Medical Fee Schedule Laws

Table 1.A.7: The Effects of Medical Fee Schedule Laws on Cash Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected Benefits Index (log)</strong></td>
<td>0.0812</td>
<td>-1.607**</td>
<td>1.903***</td>
<td>1.875***</td>
<td>1.207***</td>
</tr>
<tr>
<td></td>
<td>(0.503)</td>
<td>(0.405)</td>
<td>(0.483)</td>
<td>(0.321)</td>
<td>(0.215)</td>
</tr>
<tr>
<td><strong>Fatal Injury Rate (log)</strong></td>
<td>-0.138</td>
<td>0.0690*</td>
<td>-0.0137</td>
<td>-0.0240</td>
<td>-0.00759</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.0349)</td>
<td>(0.0451)</td>
<td>(0.0556)</td>
<td>(0.0216)</td>
</tr>
<tr>
<td><strong>Fee Schedule Laws</strong></td>
<td>0.124</td>
<td>-0.00361</td>
<td>0.0588</td>
<td>0.0390</td>
<td>-0.0353</td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td>(0.0600)</td>
<td>(0.0574)</td>
<td>(0.0579)</td>
<td>(0.0414)</td>
</tr>
</tbody>
</table>

|                              |        |        |        |        |        |
| State Fixed Effects          | X      | X      | X      | X      | X      |
| Year Fixed Effects           | X      | X      | X      |        |        |
| States without the Laws Before 1997 |        |        |        | X      |
| State-specific Time Trend    |        |        |        |        | X      |
| R-Squared                    | 0.034  | 0.184  | 0.376  | 0.360  | 0.746  |
| N                            | 1000   | 1000   | 1000   | 520    | 1000   |

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of cash benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table 1.A.8: The Effects of Medical Fee Schedule Laws on Medical Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>0.223***</td>
<td>0.135***</td>
<td>-0.00322</td>
<td>-0.0117</td>
<td>0.0195</td>
</tr>
<tr>
<td></td>
<td>(0.0761)</td>
<td>(0.0420)</td>
<td>(0.0365)</td>
<td>(0.0455)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>Fee Schedule Laws</td>
<td>0.0647</td>
<td>-0.0202</td>
<td>0.0804*</td>
<td>0.0523</td>
<td>-0.0273</td>
</tr>
<tr>
<td></td>
<td>(0.0677)</td>
<td>(0.0447)</td>
<td>(0.0444)</td>
<td>(0.0496)</td>
<td>(0.0240)</td>
</tr>
<tr>
<td>State Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>States without the Laws Before 1997</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State-specific Time Trend</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.145</td>
<td>0.048</td>
<td>0.313</td>
<td>0.293</td>
<td>0.638</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>520</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of medical benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
## 1.A.5 Physician Network Laws

Table 1.A.9: The Effects of Physician Network Laws on Cash Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Benefits Index (log)</td>
<td>0.0945</td>
<td>-1.543**</td>
<td>1.889***</td>
<td>1.893***</td>
<td>1.195***</td>
</tr>
<tr>
<td></td>
<td>(0.475)</td>
<td>(0.395)</td>
<td>(0.504)</td>
<td>(0.510)</td>
<td>(0.206)</td>
</tr>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>-0.137</td>
<td>0.0607</td>
<td>-0.0219</td>
<td>-0.0285</td>
<td>-0.0107</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
<td>(0.0382)</td>
<td>(0.0484)</td>
<td>(0.0501)</td>
<td>(0.0216)</td>
</tr>
<tr>
<td>Physician Network Laws</td>
<td>-0.242</td>
<td>-0.219</td>
<td>-0.144</td>
<td>-0.148</td>
<td>-0.192**</td>
</tr>
<tr>
<td></td>
<td>(0.236)</td>
<td>(0.153)</td>
<td>(0.162)</td>
<td>(0.163)</td>
<td>(0.0811)</td>
</tr>
</tbody>
</table>

State Fixed Effects: X X X X X
Year Fixed Effects: X X X
States without the Laws Before 1997: X X
State-specific Time Trend: X
R-Squared: 0.039 0.203 0.380 0.374 0.750
N: 1000 1000 1000 940 1000

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of cash benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table 1.A.10: The Effects of Physician Network Laws on Medical Benefit Payments

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatal Injury Rate (log)</td>
<td>0.230***</td>
<td>0.117***</td>
<td>-0.0157</td>
<td>-0.0216</td>
<td>0.0163</td>
</tr>
<tr>
<td></td>
<td>(0.0795)</td>
<td>(0.0429)</td>
<td>(0.0403)</td>
<td>(0.0416)</td>
<td>(0.0218)</td>
</tr>
<tr>
<td>Physician Network Laws</td>
<td>-0.0501</td>
<td>-0.349**</td>
<td>-0.271*</td>
<td>-0.276*</td>
<td>-0.271***</td>
</tr>
<tr>
<td></td>
<td>(0.186)</td>
<td>(0.158)</td>
<td>(0.156)</td>
<td>(0.156)</td>
<td>(0.0724)</td>
</tr>
<tr>
<td>State Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>States without the Laws Before 1997</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State-specific Time Trend</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.138</td>
<td>0.116</td>
<td>0.341</td>
<td>0.331</td>
<td>0.650</td>
</tr>
<tr>
<td>N</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>940</td>
<td>1000</td>
</tr>
</tbody>
</table>

Notes: The sample period is from 1997 to 2016. The analysis sample combines data from the National Academy of Social Insurance (NASI), Bureau of Labor Statistics (BLS), Social Security Administration (SSA), the ProPublica Website, and law volumes for the states. The outcome variable is the natural logarithm of medical benefit payments per covered worker. Standard errors are in parenthesis and clustered at state level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
2.A More Details about the Data

2.A.1 Variable Definitions

*Age*: age is measured by taking a difference between the date of birth and the date of criminal charges being filed.

*Dismiss*: Criminal charges are dismissed and defendants bear no criminal punishment. Case dismissals may occur through either judges or prosecutors.

*Bargain*: A dummy variable that is equal to 1 if charges are resolved through plea bargain, 0 if not. The data consists of various plea bargain outcomes. We infer that a case is resolved through plea bargain if one of the follows three outcomes is present: guilty, guilty to lower charges and no contest.\(^{81}\) Otherwise, we assume that defendants decided to take their cases to trials by pleading “not guilty.”

*Jury Trial*: a dummy variable indicating if the case is resolved through trial jury.

*Aggravated/Mitigated*: whether the minimum sentences defendants received come from aggravated/mitigated range under the North Carolina structured sentencing guidelines.

\(^{81}\)If plea bargain outcome is “no contest,” it means that defendants neither admit nor deny the guilt. But, the courts can determine punishment. The intent of such decisions is usually to avoid being civilly sued for essentially confessing to crime. More detail: [http://www.nolocontendere.org/differencebetweenguiltyandnocontest.html](http://www.nolocontendere.org/differencebetweenguiltyandnocontest.html)
2.A.2 Detailed Description of Data Cleaning

Our dataset comes from North Carolina Administrative Office of Courts (NCAOC). We implemented extensive cleaning based on Silveira (2017), which uses the same dataset.

2.A.2.1 Defining Criminal Cases

The data include all of criminal charges filed in North Carolina. The state court system typically define “case” with a common file number, which is frequently associated with multiple counts of charges. However, such definition of “case” does not accurately reflect the reality of criminal cases are resolved and could cause many issues in empirical analysis. First, the same file number reappears in the data if defendants violate probation conditions, which should not be counted as new criminal offenses. Second, the data suggests that defendants are often disposed of various charges associated with different file numbers on the same date. What’s more, anecdotal evidence suggests that plea bargain applies to all outstanding charges. As a result, we decided to adopt an alternative definition of “case” by using a pair of individual identifier and disposition date, which was used in Abrams and Fackler (2020).

The data cleaning begins with a total number of 2,364,028 charges which were filed in the North Carolina superior courts between 1995 and 2009. Firstly, we keep charges with meaningful outcomes, meaning either dismissed or determined by judges or juries. This procedure eliminated irrelevant cases where there are changes of venues, decisions were made state district courts, etc. The observations dropped to 1,774,336 (25% decrease). Secondly, among charges associated with the same file numbers, we keep the counts which were disposed at the earliest date, which eliminated all the charges filed due to probation violations. The number of observations dropped to 1,745,225 (1.6% decrease). Thirdly, we dropped all charges which do not consist of meaningful information on charged offense and case outcomes. The observations dropped to 1,621,345 (7.1% decrease). Fourthly, we drop the charges where defendants are convicted of misdemeanor offenses as the sentencing guidelines in Figure 2.1 does not apply to misdemeanor offenses. This step reduces the observation to 1,308,455 (19% decrease). Lastly, we dropped charges which are consolidated with other ones, which further reduced the number of observations to 911,885 (30% decrease). North Carolina state laws allow courts to consolidate less serious charges and to only issue one judgment on the basis of the most serious offense. Consolidated charges do not have influence the final sentence outcomes.

82The data includes a variable called “crdcff” and “crdfjfo,” which identifies the charge being consolidated to.
Then, we further aggregate charges at two levels in order to define criminal “cases.” First, for each file number, we keep the count associated with the most severe charge. This step reduces the number of observations to 669,576 (27% decrease). Furthermore, we implemented the same aggregation procedure for each pair of individual identifiers and disposition date. The number of observations dropped to 455,449 (32% decrease). In the end, we are able to proxy criminal “case” by the most severe charge that defendants are disposed on the same date. In developing the empirical analysis, we further restrict time frame from 2001 to 2007.

2.A.2.2 Identifying Individuals

Criminal defendants are identified based on their names and date of date of birth. Specifically, we created an identifier for each individual by using (1) the birth year and month pair and (2) the first four letter of their first and last names. Such approach helps us distinguish different individuals with the same names. What’s more, it also helps us minimize the impacts from human errors in data inputs, such as misspelling of defendants’ names and mistype of dates of birth. The individual-specific identifiers help us further defining “cases” and recidivism.

\footnote{For charges associated with the same file number, we first order them on the basis of punishment severity, which we measure by using a pair of punishment type and minimum sentence length. We keep the count(s) associated the most severe punishment. However, such procedure could not help us determine the most severe charge in some situations. For example, dismissed cases or charges with identical sentences. For those cases, we keep the charge associated with the most severe filed offense.}
2.B Additional Figures
Figure 2.B.1: Changes of Defendants’ Demographics at the *Blakely* discontinuity

Note: This figure shows the changes in defendants’ demographics at the *Blakely* discontinuity (June 2004). The curves are fitted on each side of *Blakely* with a second-degree polynomial function of the time. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of the cases charged with felony offenses in North Carolina from 2001 to 2007.
Figure 2.B.2: Cases Resolved by Plea Bargains

Note: The Figure shows the changes in the likelihood of plea bargain at the Blakely discontinuity (June 2004). The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of the cases charged with felony offenses in North Carolina from 2001 to 2007.
Figure 2.B.3: County-Level Arrest Rate in North Carolina

Note: The figure shows changes in county-level arrest rate at the *Blakely* discontinuity. Arrest rate is defined as the number of arrests made per 100,000 county population. The curves are fitted on each side of *Blakely* with a second-degree polynomial function of the time. Arrest data comes from the Uniform Crime Report (UCR).
Note: The figure shows changes in the likelihood of defendants who are charge with felony offenses but plead guilty to misdemeanors. In constructing the measures, I use the count of cases with felony charges as the denominator and the count of felony cases with defendants pleading guilty to misdemeanors as the numerator. The curves are fitted on each side of *Blakely* with a second-degree polynomial function of the time. The data consists of criminal case records from the Administrative Office of Courts (AOC) in North Carolina.
Note: The figure shows residuals from the following specification:

\[ Aggravated_{ijkc} = \beta_0 + X_i \gamma + Grid_{jk} + \theta_c + \epsilon_{ijkc} \]

The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. Each data point is the monthly average of the residuals from the regression model. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of the cases charged with felony offenses from 2001 to 2007.
Figure 2.B.6: Residual Plots of Sentence Enhancements (By Gender)

Note: The figure shows residuals from the following specification:

\[ Aggravated_{ijkc} = \beta_0 + X'_i \gamma + Grid_{jk} + \theta_c + \epsilon_{ijkc} \]

The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The control vector, \( X'_i \), excludes defendants’ gender. Each data point is the monthly average of the gender-specific residuals from the regression model. The curves are fitted on each side of Blakely with a second-degree polynomial function of the time. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample include all of the cases charged with felony offenses from 2001 to 2007.
Figure 2.B.7: Residual Plots of Sentence Enhancements (By Race)

Note: The figure shows residuals from the following specification:

\[ \text{Aggravated}_{ijkc} = \beta_0 + X_i' + Grid_{jk} + \theta_c + \epsilon_{ijkc} \]

The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The control vector, \( X_i' \), excludes defendants’ race. Each data point is the monthly average of the gender-specific residuals from the regression model. The curves are fitted on each side of Blakely with a second-degree polynomial function of the time. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample include all of the cases charged with felony offenses from 2001 to 2007.
2.C Additional Results

Table 2.C.1: Summary Statistics Pre and Post *Blakely*

<table>
<thead>
<tr>
<th></th>
<th>Pre-Blakely</th>
<th>Post-Blakely</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Male</td>
<td>0.842</td>
<td>0.364</td>
</tr>
<tr>
<td>White</td>
<td>0.394</td>
<td>0.489</td>
</tr>
<tr>
<td>Black/Hispanic</td>
<td>0.582</td>
<td>0.493</td>
</tr>
<tr>
<td>Age</td>
<td>29.543</td>
<td>10.078</td>
</tr>
<tr>
<td>Private Attorney</td>
<td>0.244</td>
<td>0.430</td>
</tr>
<tr>
<td>Public Defender</td>
<td>0.212</td>
<td>0.409</td>
</tr>
<tr>
<td>Criminal History Points</td>
<td>4.519</td>
<td>5.226</td>
</tr>
<tr>
<td>Arrest Rate (County)</td>
<td>549.183</td>
<td>233.237</td>
</tr>
<tr>
<td>Unemployment Rate (County)</td>
<td>6.423</td>
<td>1.649</td>
</tr>
<tr>
<td>Plea Bargain</td>
<td>0.869</td>
<td>0.338</td>
</tr>
<tr>
<td>Plead Guilty to Lower Offenses</td>
<td>0.249</td>
<td>0.432</td>
</tr>
<tr>
<td>Sentence Enhancement</td>
<td>0.029</td>
<td>0.167</td>
</tr>
<tr>
<td>Mitigated</td>
<td>0.074</td>
<td>0.262</td>
</tr>
<tr>
<td>Incarceration</td>
<td>0.303</td>
<td>0.460</td>
</tr>
<tr>
<td>Observations</td>
<td>103910</td>
<td>109354</td>
</tr>
</tbody>
</table>

Note: The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). Arrest rate is defined as the number of arrests made per 100,000 county population. The sample include all of the cases charged with felony offenses from 2001 to 2007.
<table>
<thead>
<tr>
<th>Offense</th>
<th>Sentence Enhancements</th>
<th>Male</th>
<th>Black/Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assault</td>
<td>0.029</td>
<td>0.883</td>
<td>0.618</td>
</tr>
<tr>
<td>Burglary</td>
<td>0.021</td>
<td>0.919</td>
<td>0.461</td>
</tr>
<tr>
<td>Drugs</td>
<td>0.015</td>
<td>0.855</td>
<td>0.686</td>
</tr>
<tr>
<td>Embezzlement</td>
<td>0.017</td>
<td>0.406</td>
<td>0.374</td>
</tr>
<tr>
<td>Forgery</td>
<td>0.013</td>
<td>0.557</td>
<td>0.466</td>
</tr>
<tr>
<td>Fraud</td>
<td>0.016</td>
<td>0.584</td>
<td>0.448</td>
</tr>
<tr>
<td>Homicide</td>
<td>0.102</td>
<td>0.893</td>
<td>0.610</td>
</tr>
<tr>
<td>Kidnapping</td>
<td>0.048</td>
<td>0.934</td>
<td>0.606</td>
</tr>
<tr>
<td>Larceny</td>
<td>0.012</td>
<td>0.836</td>
<td>0.439</td>
</tr>
<tr>
<td>Other</td>
<td>0.017</td>
<td>0.777</td>
<td>0.556</td>
</tr>
<tr>
<td>Robbery</td>
<td>0.038</td>
<td>0.927</td>
<td>0.741</td>
</tr>
<tr>
<td>Sexual Assault</td>
<td>0.039</td>
<td>0.974</td>
<td>0.462</td>
</tr>
<tr>
<td>Traffic</td>
<td>0.017</td>
<td>0.920</td>
<td>0.548</td>
</tr>
<tr>
<td>Weapon Offense</td>
<td>0.017</td>
<td>0.970</td>
<td>0.718</td>
</tr>
<tr>
<td>Total</td>
<td>0.021</td>
<td>0.840</td>
<td>0.575</td>
</tr>
<tr>
<td>Observations</td>
<td>213264</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). Offense refers to the type of crimes defendants are charged with.
Table 2.C.3: The Effects of *Blakely* on Sentence Enhancements (Different Bandwidths)

<table>
<thead>
<tr>
<th>Bandwidths</th>
<th>(1) 6 Months</th>
<th>(2) 12 Months</th>
<th>(3) 24 Months</th>
<th>(4) Whole Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>-0.0240***</td>
<td>-0.0206***</td>
<td>-0.0176***</td>
<td>-0.0185***</td>
</tr>
<tr>
<td></td>
<td>(0.00197)</td>
<td>(0.00118)</td>
<td>(0.00120)</td>
<td>(0.00127)</td>
</tr>
<tr>
<td>N</td>
<td>28051</td>
<td>54704</td>
<td>108257</td>
<td>186825</td>
</tr>
</tbody>
</table>

Note: The pre-*Blakely* mean of sentence enhancements is 0.029. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. We perform local linear regression in all of the specifications. Each specification includes a rich set of demographic controls, county fixed effects, month of year fixed effects and guidelines fixed effects. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The standard errors are clustered at month level. Significance Levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table 2.C.4: The Effects of *Blakely* on Gender Disparity in Sentence Enhancements (Different Bandwidths)

<table>
<thead>
<tr>
<th>Bandwidths</th>
<th>(1) 6 Months</th>
<th>(2) 12 Months</th>
<th>(3) 24 Months</th>
<th>(4) Whole Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>-0.0276</td>
<td>-0.0158**</td>
<td>-0.0117***</td>
<td>-0.0117***</td>
</tr>
<tr>
<td></td>
<td>(0.0286)</td>
<td>(0.00229)</td>
<td>(0.00250)</td>
<td>(0.00196)</td>
</tr>
<tr>
<td>Male</td>
<td>0.00749***</td>
<td>0.00534***</td>
<td>0.00704***</td>
<td>0.00705***</td>
</tr>
<tr>
<td></td>
<td>(0.00146)</td>
<td>(0.00189)</td>
<td>(0.00186)</td>
<td>(0.00147)</td>
</tr>
<tr>
<td>1[Blakely] * Male</td>
<td>-0.00765*</td>
<td>-0.00580**</td>
<td>-0.00703***</td>
<td>-0.00807***</td>
</tr>
<tr>
<td></td>
<td>(0.00369)</td>
<td>(0.00270)</td>
<td>(0.00232)</td>
<td>(0.00175)</td>
</tr>
<tr>
<td>N</td>
<td>28051</td>
<td>54704</td>
<td>108257</td>
<td>186825</td>
</tr>
</tbody>
</table>

Note: The pre-*Blakely* mean of sentence enhancements is 0.029. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. We perform local linear regression in all of the specifications. Each specification includes a rich set of demographic controls, county fixed effects, month of year fixed effects and guidelines fixed effects. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The standard errors are clustered at month level. Significance Levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table 2.C.5: The Effects of *Blakely* on Gender Disparity in Sentence Enhancements (Different Bandwidths)

<table>
<thead>
<tr>
<th>Bandwidths</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>-0.0347</td>
<td>-0.0206***</td>
<td>-0.0182***</td>
<td>-0.0181***</td>
</tr>
<tr>
<td></td>
<td>(0.0275)</td>
<td>(0.00146)</td>
<td>(0.00184)</td>
<td>(0.00159)</td>
</tr>
<tr>
<td>Black/Hispanic</td>
<td>-0.00162</td>
<td>-0.00103</td>
<td>-0.00287*</td>
<td>-0.00147</td>
</tr>
<tr>
<td></td>
<td>(0.00259)</td>
<td>(0.00205)</td>
<td>(0.00167)</td>
<td>(0.00124)</td>
</tr>
<tr>
<td>1[Blakely] * Black/Hispanic</td>
<td>0.000956</td>
<td>-0.0000163</td>
<td>0.00106</td>
<td>-0.000763</td>
</tr>
<tr>
<td></td>
<td>(0.00305)</td>
<td>(0.00241)</td>
<td>(0.00189)</td>
<td>(0.00140)</td>
</tr>
<tr>
<td>N</td>
<td>28051</td>
<td>54704</td>
<td>108257</td>
<td>186825</td>
</tr>
</tbody>
</table>

Note: The pre-*Blakely* mean of sentence enhancements is 0.029. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. We perform local linear regression in all of the specifications. Each specification includes a rich set of demographic controls, county fixed effects, month of year fixed effects and guidelines fixed effects. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The standard errors are clustered at month level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
### Table 2.C.6: The Effects of Blakely on Plea Bargain

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>0.00176</td>
<td>0.000415</td>
<td>0.000461</td>
<td>0.000270</td>
<td>0.00137</td>
</tr>
<tr>
<td></td>
<td>(0.00405)</td>
<td>(0.00160)</td>
<td>(0.00170)</td>
<td>(0.00169)</td>
<td>(0.00237)</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Month-of-Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Guidelines Fixed Effects</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local Linear Regression</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Polynomial Regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>213264</td>
<td>186825</td>
<td>186825</td>
<td>186825</td>
<td>186825</td>
</tr>
</tbody>
</table>

Note: The pre-Blakely mean of plea bargain is 86.9%. The outcome variable is a dummy indicating whether a case is resolved through plea bargain. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of criminal cases charged with felony offenses from 2001 to 2007. The standard errors are clustered at month level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
Table 2.C.7: The Effects of *Blakely* on Sentence Enhancements (Plea Bargain Sub-sample)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1[Blakely]</td>
<td>-0.0179***</td>
<td>-0.0179***</td>
<td>-0.0186***</td>
<td>-0.0160***</td>
<td>-0.0185***</td>
</tr>
<tr>
<td></td>
<td>(0.00126)</td>
<td>(0.00126)</td>
<td>(0.00123)</td>
<td>(0.00188)</td>
<td>(0.00124)</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Month-of-Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guidelines Fixed Effects</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local Linear Regression</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Polynomial Regression</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>183894</td>
<td>182782</td>
<td>182782</td>
<td>182782</td>
<td>182782</td>
</tr>
</tbody>
</table>

Note: The pre-*Blakely* mean of sentence enhancements is 0.031. The outcome variable is a dummy indicating whether a defendant receives a sentence enhancement. The analysis sample combines data from the following sources: criminal case records from the Administrative Office of Courts (AOC) in North Carolina, county-level arrest data from the Uniform Crime Report (UCR), and unemployment data from the Bureau of Labor Statistics (BLS). The sample includes all of criminal cases that were charged with felony offenses and resolved through plea bargain from 2001 to 2007. The standard errors are clustered at month level. Significance Levels: * p < 0.1, ** p < 0.05, *** p < 0.01
References


F Burton and Emily Spieler. Workers’ compensation and older workers. 2001.


Jonah B Gelbach. Locking the door to discovery-assessing the effects of twombly and iqbal on access to discovery. Yale LJ, 121:2270, 2011.


