

# The Effects of Voluntary and Presumptive Sentencing Guidelines

Griffin Edwards,<sup>\*</sup> Stephen Rushin,<sup>\*\*</sup> and Joseph Colquitt<sup>\*\*\*</sup>

*This Article empirically illustrates that the introduction of voluntary and presumptive sentencing guidelines at the state level can contribute to statistically significant reductions in sentence length, interjudge disparities, and racial disparities.*

*For much of American history, judges had largely unguided discretion to select criminal sentences within statutorily authorized ranges. But in the mid- to late-twentieth century, states and the federal government began experimenting with sentencing guidelines designed to rein in judicial discretion to ensure that similarly situated offenders received comparable sentences. Some states have made their guidelines voluntary, while others have made their guidelines presumptive or mandatory, meaning that judges must generally adhere to them unless they can justify a departure.*

*In order to explore the effects of both voluntary and presumptive sentencing guidelines on judicial behavior, this Article relies on a comprehensive data set of 221,934 criminal sentences handed down by 355 different judges in Alabama between 2002 and 2015. This data set provides a unique opportunity to address this empirical question, in part because of Alabama's legislative history. Between 2002 and 2006, Alabama had no sentencing guidelines. In 2006, the state introduced voluntary sentencing guidelines. Then in 2013, the state made these sentencing guidelines presumptive for some nonviolent offenses.*

*Using a difference-in-differences framework, we find that the introduction of voluntary sentencing guidelines in Alabama coincided with a decrease in average sentence length of around seven months. When the same guidelines became presumptive, the average sentence length dropped by almost two years. Further, using a triple-difference framework, we show that the adoption of these sentencing guidelines coincided with around eight- to twelve-month reductions in race-based sentencing disparities and substantial reductions in interjudge sentencing disparities across all classes of offenders. Combined, these data*

---

<sup>\*</sup> Griffin Edwards is an Associate Professor at the University of Alabama, Birmingham. Ph.D., Emory University.

<sup>\*\*</sup> Stephen Rushin is an Associate Professor at Loyola University Chicago School of Law. Ph.D. and J.D., University of California, Berkeley.

<sup>\*\*\*</sup> Joseph Colquitt is Professor Emeritus of Law at the University of Alabama School of Law and Chairman of the Alabama Sentencing Commission. J.D., University of Alabama. From 1971 to 1991, Professor Colquitt served as circuit judge on the Sixth Circuit of Alabama. He also served four terms as presiding judge. We thank our colleagues who provided valuable feedback on earlier drafts of this paper.

*suggest that voluntary and presumptive sentencing guidelines can help states combat inequality in their criminal justice systems while controlling the sizes of their prison populations.*

INTRODUCTION.....	2
I. EXISTING LITERATURE ON STATE SENTENCING GUIDELINES .....	8
A. Constitutional Limits on Sentencing Guidelines .....	9
B. Predictions About <i>Blakely</i> -Compliant State Sentencing Guidelines .....	16
C. Existing Research.....	18
II. METHODOLOGY .....	22
A. Background on Alabama’s Sentencing Guidelines.....	22
B. Data Set.....	26
C. Models .....	28
1. <i>Measuring the Effect on Sentence Length</i> .....	28
2. <i>Measuring the Effect on Racial Sentence Disparities</i> .....	30
3. <i>Measuring the Effect on Interjudge Sentencing             Disparities</i> .....	31
III. FINDINGS.....	32
A. Trends in Raw Data .....	33
B. Reduction in Sentence Lengths.....	41
C. Reductions in Racial Disparities .....	47
D. Reductions in Interjudge Disparities.....	52
E. Limitations of Study .....	56
IV. IMPLICATIONS .....	58
A. Importance of Presumptive Sentencing Guidelines .....	59
B. The Costs of <i>Blakely</i> .....	60
C. Alabama as a Blueprint.....	61
CONCLUSION .....	64

## Introduction

For much of American history, judges had largely unregulated discretion to issue sentences within statutory limits.<sup>1</sup> By the mid-twentieth century, a strong body of research illustrated that this unfettered discretion contributed to similarly situated criminal offenders receiving widely

---

1. For example, before the federal government adopted the Federal Sentencing Guidelines, Professor Crystal Yang recounted an observation by Judge Marvin Frankel of the Southern District of New York, who noted that “the almost wholly unchecked and sweeping powers we give to judges in the fashioning of sentences are terrifying and intolerable for a society that professes devotion to the rule of law.” Crystal S. Yang, *Have Interjudge Sentencing Disparities Increased in an Advisory Guidelines Regime? Evidence from Booker*, 89 N.Y.U. L. REV. 1268, 1270 (2014).

disparate sentences.<sup>2</sup> In response, states began experimenting<sup>3</sup> with sentencing guidelines in hopes of bringing “uniformity and proportionality to sentencing, meaning that defendants with similar criminal histories who committed similar crimes would receive similar sentences.”<sup>4</sup> Generally, these guidelines use factors like the severity of the offense and the offender’s criminal history to recommend or require that a judge issue a specific sentence length within a statutory range.<sup>5</sup> While the U.S. Supreme Court in *Blakely v. Washington*<sup>6</sup> established some limits on the ability of states to employ these sorts of criminal sentencing guidelines,<sup>7</sup> a significant number of states use some type of guidelines to regulate judicial behavior at sentencing.<sup>8</sup> Scholars generally sort sentencing guidelines into two different

---

2. See *infra* subpart I(A) and notes 42–47 (describing a series of prior studies showing that unfettered sentencing discretion without applicable guidelines contributed to widespread disparities).

3. Kevin R. Reitz, *The New Sentencing Conundrum: Policy and Constitutional Law at Cross-Purposes*, 105 COLUM. L. REV. 1082, 1085 (2005) (describing states as “laboratories of innovation” in sentencing reform).

4. Kelly Lyn Mitchell, *State Sentencing Guidelines: A Garden Full of Variety*, FED. PROBATION, Sept. 2017, at 28, 29.

5. *Id.* (noting that “[t]he two primary determinants of the sentence under sentencing guidelines systems are offense severity and criminal history” with most jurisdictions using either a grid or a worksheet system to calculate points based on these variables).

6. 542 U.S. 296 (2004).

7. By many accounts, *Blakely v. Washington* was a watershed moment in the history of criminal procedure and sentencing. *Id.* at 301–05 (finding that the Sixth Amendment jury trial guarantee requires states to submit any factual finding that may increase the maximum allowable punishment to a jury to be decided beyond a reasonable doubt). For much of American history, and “even in the heyday of the Warren Court,” federal constitutional law had little to say about procedures used in criminal sentencing. Reitz, *supra* note 3, at 1083. As one scholar put it, the Constitution generally provided “no meaningful constitutional brake on the nation’s thirty-year revolution in the use of prisons, jails, and community sanctions.” *Id.* at 1084. Courts generally deferred to state choices as to the severity of criminal punishment and the procedures used to arrive at a particular sentence. *Id.* But in 2004, the U.S. Supreme Court delivered one of “the most audacious” criminal procedure rulings in history, which some scholars compared to a “legal earthquake, a forty-car pileup, a bombshell, . . . a bull in a china shop” and perhaps “the most significant constitutional decision in criminal justice since *Miranda*.” *Id.* at 1086. It held that the Washington sentencing guidelines unconstitutionally denied criminal defendants their Sixth Amendment jury trial guarantee; any factual finding that increases the maximum punishment for a criminal defendant must be submitted to a jury and decided beyond a reasonable doubt. *Blakely*, 542 U.S. at 301–05 (describing the extension of *Apprendi v. New Jersey*, 530 U.S. 466 (2000), to the facts in *Blakely*). This ruling effectively invalidated presumptive sentencing guidelines in states across the country, including Alaska, Arizona, California, Colorado, Indiana, Minnesota, New Jersey, New Mexico, North Carolina, Ohio, Oregon, Tennessee, and Washington. John F. Pfaff, *The Continued Vitality of Structured Sentencing Following Blakely: The Effectiveness of Voluntary Guidelines*, 54 UCLA L. REV. 235, 250–51 (2006) (“Thus, twelve states besides Washington currently employ sentencing regimes that likely run afoul of *Blakely*.”). In her dissent to the *Blakely* decision, Justice Sandra Day O’Connor wrote that what she “feared most has now come to pass: Over 20 years of sentencing reform are all but lost.” *Id.* at 326 (O’Connor, J., dissenting).

8. It is worth noting that before *Blakely*, a large number of states began using guidelines, with many of them binding or presumptive guidelines that limited the ability of judges to stray outside a narrow guideline range in issuing criminal sentences. Rodney L. Engen, *Assessing Determinate and Presumptive Sentencing—Making Research Relevant*, 8 CRIMINOLOGY & PUB. POL’Y 323, 323

categories: voluntary sentencing guidelines and presumptive sentencing guidelines.

Voluntary guidelines are “a starting point or suggestion for sentencing,” while presumptive or mandatory guidelines “connote[] that the sentences established by the guidelines are required.”<sup>9</sup> According to one estimate, eight states (Arkansas, Delaware, Maryland, Massachusetts, Michigan, Pennsylvania, Utah, and Virginia) and the District of Columbia currently use voluntary sentencing guidelines.<sup>10</sup> By contrast, five states (Kansas, Minnesota, North Carolina, Oregon, and Washington) use presumptive or mandatory sentencing guidelines.<sup>11</sup> Only one state—Alabama—uses a combination of both voluntary and presumptive sentencing guidelines.<sup>12</sup> The majority of states employ no formal sentencing guidelines.<sup>13</sup>

This wide variation in sentencing policy raises several important empirical questions. How well do these state sentencing guidelines address interjudge disparities and racial disparities relative to states without guidelines? Have states with presumptive sentencing guidelines better addressed disparities than those with voluntary guidelines? And how has the implementation of sentencing guidelines influenced overall sentence lengths? Unfortunately, the available literature on this topic is relatively sparse—particularly legal scholarship on the comparative usefulness of different approaches to sentencing guidelines at the state level. While several studies have explored the effects of the Federal Sentencing Guidelines,<sup>14</sup> a relatively small body of literature has analyzed the effect of different state sentencing-guideline policies on judicial behavior.<sup>15</sup> As one scholar noted, studies “comparing sentencing practices with and without guidelines in place [are] *exceptionally* rare.”<sup>16</sup> And the effect of different forms of sentencing guidelines on sentence length may be “one of the most important questions” that researchers “seldom address.”<sup>17</sup>

This Article makes a significant addition to the existing empirical literature by turning much-needed attention to state sentencing guidelines,

---

(2009) (finding that around seventeen states had adopted mandatory or binding sentencing guidelines, while another eighteen had adopted some sort of presumptive sentencing guidelines).

9. Mitchell, *supra* note 4, at 34.

10. *Id.* at 36 tbl.5.

11. *Id.*

12. *Id.* (noting that “Alabama has two sets of guidelines; only the presumptive guidelines . . . would be characterized as mandatory”).

13. *See id.* (showing that only fourteen states have either advisory or mandatory sentencing guidelines, with the remaining states not employing such guidelines).

14. *See infra* notes 103–11 and accompanying text (describing a handful of the empirical studies on federal sentencing guidelines).

15. *See infra* notes 114–19 and accompanying text (describing the smaller body of literature on this particular topic).

16. Engen, *supra* note 8, at 324.

17. *Id.* at 329.

which affect substantially more criminal defendants than their federal counterparts.<sup>18</sup> We empirically illustrate that the introduction of both voluntary and presumptive sentencing guidelines at the state level can contribute to statistically significant reductions in sentence length, interjudge disparities, and racial disparities. To do this, we focus our research on a series of legal events in Alabama that allow us to test the effects of both voluntary and presumptive sentencing guidelines. Alabama is unique among American states in its recent experimentation with different kinds of sentencing guidelines. For much of its history, Alabama used no sentencing guidelines.<sup>19</sup> While state criminal statutes established applicable sentence ranges, trial judges were free to issue whatever sentence lengths they felt appropriate within the confines of these statutory ranges. Then in 2006, Alabama introduced voluntary sentencing guidelines for some criminal offenses, particularly personal offenses (like murder, rape, assault, robbery, and manslaughter), burglary offenses, nonviolent drug offenses, and nonviolent property offenses (like theft and forgery).<sup>20</sup> In these cases, Alabama law required judges to complete a sentencing worksheet that recommended a sentence length based on a consideration of a number of factors relevant to the offender's culpability.<sup>21</sup> But judges were free to deviate from these guidelines without penalty, and such deviations from the guidelines were generally unappealable.<sup>22</sup> This meant that, between 2006 and 2013, Alabama employed voluntary sentencing guidelines for some offenses and no sentencing guidelines for other, so-called nonworksheet offenses.

Then in 2013, “the Alabama Legislature changed the Standards for non-violent offenses . . . from voluntary to presumptive recommendations.”<sup>23</sup>

---

18. DANIELLE KAEBLE & MARY COWHIG, CORRECTIONAL POPULATIONS IN THE UNITED STATES, 2016, BUREAU OF JUSTICE STATISTICS 11 app. tbl.1 (2018), <https://www.bjs.gov/content/pub/pdf/cpus16.pdf> [<https://perma.cc/4NVT-Y2KE>] (showing that in 2016 the federal government incarcerated around 320,000 inmates compared to the approximately 6,262,000 inmates incarcerated by state systems).

19. ALA. SENTENCING COMM'N, COLLABORATIVE SUCCESS: ALABAMA IMPLEMENTS SENTENCING STANDARDS 5–7 (2007), <http://sentencingcommission.alacourt.gov/media/1045/2007-annual-report.pdf> [<https://perma.cc/KLX2-WQ8X>] (describing the establishment of the Alabama Sentencing Commission, in part because of a recognition that the prior system without guidelines resulted in unwarranted sentencing disparities and overcrowding of prisons and jails).

20. *Id.* at 11 (“The initial voluntary sentencing standards were adopted and became effective October 1, 2006,” as part of Act No. 2006-312). It is worth mentioning, as well, that the Sentencing Commission established the sentencing guidelines after it developed reliable data from five years of felony data, which allowed it to estimate the effects of various sentencing standards on overall system-wide outcomes. *Id.* at 7–8.

21. *Id.* at 8 (describing the use of physical and electronic worksheets for “standards implementation”).

22. *Id.* at 11 (explaining that the guidelines were “voluntary, non-appealable, historically based, time imposed sentencing recommendations developed for 26 felony offenses”).

23. ALA. SENTENCING COMM'N, PRESUMPTIVE AND VOLUNTARY SENTENCING STANDARDS: MANUAL 15 (2016) [hereinafter 2016 MANUAL], <http://sentencingcommission.alacourt.gov/media/1065/2016-presumptive-manual.pdf> [<https://perma.cc/2YHJ-NUX3>].

This meant that, for many nonviolent offenses, judges in Alabama were now legally required to issue sentence lengths consistent with the sentencing worksheets developed by the Alabama Sentencing Commission. Judges could deviate from these presumptive sentence lengths, but only if they could identify an applicable aggravating or mitigating circumstance.<sup>24</sup> Effectively, this meant that between 2013 and 2015, Alabama used presumptive sentencing guidelines for many nonviolent offenses, voluntary guidelines for certain violent offenses, and no guidelines for nonworksheet offenses. This series of legal events in a single jurisdiction over a relatively short period of time provides an opportunity to explore the effects of voluntary and presumptive sentencing guidelines on judicial behavior.

To do so, this Article takes advantage of a comprehensive data set of all 221,934 criminal sentences handed down by 355 different judges in Alabama between 2002 and 2015. Our data set, which we obtained directly from the Alabama Sentencing Commission, includes details on the demographic profile, criminal history, and other important variables on each criminal defendant.<sup>25</sup> It also allows us to track the behavior of judges over this fourteen-year time frame to see how their sentences changed in response to these various sentencing procedures.<sup>26</sup> Using a difference-in-differences framework, we find that the introduction of voluntary sentencing guidelines in Alabama coincided with a decrease in average sentence length of around seven months, with the effect being most statistically significant among judges in the middle quartiles of sentencing “toughness.” When the same guidelines became presumptive, the average sentence length dropped by almost two years, with the effect being most significant among judges who previously issued the strictest sentences. Further, using a triple-difference framework, we show that the adoption of these sentencing guidelines coincided with around eight- to twelve-month reductions in race-based sentencing disparities and substantial reductions in interjudge sentencing disparities across all classes of offenders.

Overall, our data contribute to the empirical literature on the effects of sentencing guidelines on judicial behavior. For one thing, our data suggest that both voluntary and presumptive sentencing guidelines were at least somewhat effective at influencing judicial sentencing behavior in Alabama. We find that voluntary sentencing guidelines may have contributed to some reductions in sentencing disparities and lower overall sentence lengths,

---

24. *Id.* at 16–17, 19, 28–30 (describing the operation of the sentencing guidelines and then describing the aggravating and mitigating circumstances developed by the Sentencing Commission).

25. *See infra* subpart II(B) (describing the data set, including all variables we considered).

26. *See infra* subpart II(B).

consistent with the stated goals of the Alabama Sentencing Commission.<sup>27</sup> However, our data suggest that presumptive sentencing guidelines are comparatively more effective at altering judicial behavior.<sup>28</sup> And perhaps most importantly, presumptive guidelines in Alabama were particularly more effective than voluntary guidelines at influencing the behavior of judges who were, a priori, more punitive than their peers.<sup>29</sup> In the wake of the U.S. Supreme Court's decision in *Blakely v. Washington*,<sup>30</sup> it appears that a number of states may have moved from presumptive to voluntary guidelines.<sup>31</sup> Our data suggest that such a shift in sentencing-guideline structure may contribute to additional disparities—particularly among judges at the top and bottom of comparative punitiveness. These results are also consistent with a number of studies that have found that the Court's decision in *United States v. Booker*<sup>32</sup> similarly contributed to a rise in interjudge and racial disparities in the federal system.<sup>33</sup> Finally, given that many states have not yet adopted any formal sentencing guidelines,<sup>34</sup> these findings should make other states consider changes to their sentencing laws. Our data suggest that many of these states could benefit from the adoption of sentencing guidelines similar to those of Alabama in order to combat inequality in their criminal justice systems while controlling the sizes of their prison populations.

This Article proceeds in four parts. Part I examines the existing literature on state sentencing guidelines. It walks through the history of sentencing guidelines in the United States, including how the Court has acted to regulate the use of these guidelines at the state level. It also details some of the existing research on the effects of state sentencing guidelines. Part II explains our methodology. Part III summarizes our results, and Part IV considers some of the implications of our study.

---

27. See *infra* subparts III(A)–(D) (showing that voluntary guidelines may have contributed to modest reductions in sentence length and modest reductions in disparity, although some are not statistically significant).

28. See *infra* subparts III(A)–(D) (showing that presumptive guidelines contributed to relatively large and statistically significant changes in sentences and reductions in sentence disparities, with the effect being most salient among the judges with a history of being most punitive).

29. See *infra* subparts III(A)–(D).

30. 542 U.S. 296 (2004) (holding that Washington state could not require trial judges to make factual findings based on a standard less than beyond a reasonable doubt during the penalty phase that then increase the maximum punishment for a criminal offender).

31. For more discussion of this point, see *infra* notes 170, 172, and accompanying text (describing the number of states that had presumptive or mandatory sentencing guidelines before *Blakely* and the number that utilize similar systems today).

32. 543 U.S. 220 (2005) (holding that the Federal Sentencing Guidelines impermissibly gave trial judges the authority to make factual findings by a standard lower than beyond a reasonable doubt and use these determinations to raise the overall maximum sentence length in criminal cases in violation of the Sixth Amendment).

33. See *infra* notes 110–13 and accompanying text (describing studies like that by Professor Yang that show the increase in racial disparities and interjudge disparities after *Booker*).

34. Mitchell, *supra* note 4, at 36 tbl.5.

### I. Existing Literature on State Sentencing Guidelines

For much of American history, jurisdictions empowered trial judges or correctional officials to determine the appropriate length of punishment for criminal offenders within statutorily authorized ranges.<sup>35</sup> So, imagine a state that statutorily permitted imprisonment for between five and twenty years for an offender convicted of a second-degree felony. That state might permit the trial court judge to exercise discretion in assigning a punishment of between five- and twenty-years imprisonment. Or conversely, that state might allow a state department of corrections to imprison an offender convicted of a second-degree felony for between five and twenty years, depending on whether the offender has been properly rehabilitated.

In either case, such a system introduces significant discretion into the process of criminal punishment. Those holding such discretion—be they trial court judges, correctional officials, or parole boards—should theoretically use it to tailor criminal sentences to the specific culpability of each criminal offender. But officials can also abuse this discretion. Some judges or correctional officials may be more punitive than others, leading to variation in the actual sentences served by similarly situated offenders. And some judges or correctional officials may exercise their discretion in a way that treats offenders differently based on impermissible reasons, such as the offender's race or gender. It should come as no surprise, then, that previous research found that discretionary, indeterminate, and unguided sentencing procedures contributed to disparities in the criminal justice system.<sup>36</sup>

Around the mid- to late-twentieth century, a number of American jurisdictions began experimenting with determinate sentences and sentencing guidelines designed to regulate the exercise of discretion in criminal punishment.<sup>37</sup> These reforms sought “to eliminate both actual and perceived disparities in criminal sentencing by making it more regular and more transparent.”<sup>38</sup> While these sentencing guidelines took many forms, they rapidly spread to jurisdictions throughout the country.

But soon thereafter, the U.S. Supreme Court radically reshaped the world of sentencing guidelines with a series of major cases that held that some of these sentencing guidelines—particularly those that mandatorily or presumptively limited judicial discretion—may be unconstitutional under the Sixth Amendment. Thus, in the years that have followed, states and the federal government have had to alter their sentencing guidelines to comply with this new understanding of the Constitution. Some feared that efforts to

---

35. See *infra* subpart I(A).

36. See *infra* note 42 and accompanying text (describing prior studies that illustrated the inconsistency in punishments before states and the federal government began implementing sentencing guidelines).

37. See *infra* subpart I(A).

38. Brief of Alabama et al. as Amici Curiae in Support of Respondent at 1, *Blakely v. Washington*, 542 U.S. 296 (2004) (No. 02-1632).



comply with the procedural requirements demanded by the Court might exact a high cost on already strapped state budgets, potentially resulting in some jurisdictions moving away from the use of sentencing guidelines. The subparts that follow walk through the history of sentencing guidelines, the Supreme Court's regulation of sentencing guidelines, and the existing empirical literature on the effect of sentencing guidelines on judicial behavior.

#### A. *Constitutional Limits on Sentencing Guidelines*

During the first half of the twentieth century, “judges [often] had broad discretion to choose a sentence within a large statutory range, and then the length of time a serious felon actually spent in prison often depended on a later decision by a parole board as to whether his release comported with public safety.”<sup>39</sup> During this period, as one commentator put it, “it [would be] difficult to say that there was much of a ‘law’ of sentencing.”<sup>40</sup> It was effectively a “Wild West of unregulated discretion.”<sup>41</sup> In theory, this discretionary system gave judges the power to individualize punishment to reflect the moral blameworthiness of each offender. After all, no legislature could possibly create a list of all relevant factors that a judge should consider when assessing the moral blameworthiness of a criminal offender. Thus, by giving the trial judge complete discretion to consider virtually any factor relevant to an offender's moral blameworthiness, discretionary sentencing policies should have contributed to highly individualized sentences within statutory ranges.

But, early empirical studies of these discretionary systems suggested that they contributed to significant disparities.<sup>42</sup> These studies relied on a number of different methodologies. Some looked at courts that randomly assigned cases to trial judges and compared the resulting sentences given by each judge, under the assumption that random assignment of cases would result in a similar caseload for each judge.<sup>43</sup> Other studies attempted to

---

39. JOHN KAPLAN ET AL., *CRIMINAL LAW: CASES AND MATERIALS* 97 (8th ed. 2017).

40. Steven L. Chanenson, *The Next Era of Sentencing Reform*, 54 *EMORY L.J.* 377, 392 (2005).

41. *Id.*

42. *See, e.g.*, Shari Seidman Diamond & Hans Zeisel, *Sentencing Councils: A Study of Sentence Disparity and Its Reduction*, 43 *U. CHI. L. REV.* 109, 111–16 (1975) (summarizing many of the research projects exploring disparities in discretionary sentences handed down by trial judges). In 1975, Professors Hans Zeisel and Shari Seidman Diamond found that researchers proved the existence of disparities in criminal sentences in three ways: through comparing the severity of sentences made by trial judges who had cases randomly assigned to them, by exploring differences between sentences handed down on factually similar cases, and through having judges issue simulated sentences under procedural circumstances that gave them substantial discretion. All three methodologies revealed that trial judges will exercise discretion in a manner that contributes to disparities in sentence length. *Id.*

43. *See, e.g.*, George Everson, *The Human Element in Justice*, 10 *J. CRIM. L. & CRIMINOLOGY* 90, 95 (1919) (comparing the sentences given by forty-two magistrate judges in New York City where cases were randomly assigned to judges and finding that the frequency of suspended

analyze sentence disparities for offenders convicted for the same crimes under factually similar circumstances.<sup>44</sup> And still other researchers conducted lab experiments that assigned factually identical cases to different judges and measured any resulting disparities.<sup>45</sup> Virtually all of these studies found that discretionary sentencing models contributed to widespread disparities in sentences. As one scholar put it, discretionary policies result in “gross disparity in sentencing, with different sentences imposed upon similar offenders who have committed similar offenses by the same judge on different days, different judges on different days, different judges on the same day, and different judges in different jurisdictions.”<sup>46</sup> These studies also found that an offender’s race frequently played a significant role in his sentence.<sup>47</sup>

In part because of this empirical evidence of disparities in the criminal justice system, states and the federal government soon began to experiment with guidelines to regulate judicial discretion.<sup>48</sup> Minnesota and Pennsylvania were on the forefront in developing guidelines in the 1970s designed to limit judicial discretion in sentencing.<sup>49</sup> In the years that followed, these guideline systems spread rapidly throughout the country. By 2002, one estimate found that around seventeen states had adopted mandatory sentencing guidelines, which required trial judges to consider various aggravating or mitigating factors and assign a punishment according to a specified formula or worksheet.<sup>50</sup> Another eighteen states employed presumptive sentencing guidelines, which required trial judges to adhere to such a formula or worksheet in assigning criminal punishment but allowed trial judges to depart from these recommendations in extraordinary circumstances.<sup>51</sup> Around eight

---

sentences for public intoxication varied from less than 1% to as high as 83% depending on the judge assigned to the case); Frederick J. Gaudet et al., *Individual Differences in the Sentencing Tendencies of Judges*, 23 J. CRIM. L. & CRIMINOLOGY 811, 813–14 (1933) (finding that, when comparing judges randomly assigned criminal cases, the frequency and severity of sentences vary).

44. Diamond & Zeisel, *supra* note 42, at 112–14 (describing these studies and their respective limitations).

45. *Id.* at 114–16 (describing a number of studies following this methodological approach, generally resulting in a finding of significant disparities across judges).

46. Richard Singer, *In Favor of “Presumptive Sentences” Set by a Sentencing Commission*, 24 CRIME & DELINQ. 401, 402 (1978).

47. *See, e.g.*, Lawrence P. Tiffany et al., *A Statistical Analysis of Sentencing in Federal Courts: Defendants Convicted After Trial, 1967-1968*, 4 J. LEGAL STUD. 369, 387–88 (1975) (finding that equally situated black defendants in federal court received significantly longer sentences than white defendants).

48. Chanenson, *supra* note 40, at 395 (“Ushered in by the attacks on unguided sentencing systems, legislative, judicial, and administrative bodies started to experiment with guidance for sentencing decisions.”).

49. *Id.* at 396 (“By the late-1970s, states took the lead on presumptive sentencing guidelines with Minnesota and Pennsylvania in the vanguard.”).

50. Engen, *supra* note 8, at 323 (citing DON STEMEN ET AL., OF FRAGMENTATION AND FERMENT, 1975–2002 (2005)).

51. *Id.*

states made their sentencing guidelines merely voluntary on trial judges, meaning that trial judges were encouraged but not required to follow the operative worksheet or formula.<sup>52</sup> The remaining states appeared to have no formal sentencing guidelines in criminal cases, meaning that judges had largely unguided discretion to assign any punishment within the statutorily authorized range.<sup>53</sup> The federal government also adopted the Federal Sentencing Guidelines in 1984.<sup>54</sup> Many of these sentencing guidelines, including the Federal Sentencing Guidelines, gave the trial judges the responsibility of making factual findings about the case at the sentencing stage by a preponderance of evidence. These factual findings would then force the judge to issue a criminal sentence within a prespecified range.

By the end of the twentieth century, it appeared as if mandatory and presumptive sentencing guidelines based on factual findings by a trial judge at the sentencing phase of a trial were rapidly becoming the norm. But things changed dramatically between 1999 and 2005, when the U.S. Supreme Court issued a series of holdings that effectively invalidated many of these sentencing guidelines, forcing many jurisdictions back to the drawing board. The first preview of the Supreme Court's future jurisprudence in this field came in 1999 in a footnote in *Jones v. United States*.<sup>55</sup> That case dealt with whether a federal carjacking statute constituted three separate criminal offenses or one single criminal offense.<sup>56</sup> In the sixth footnote to the opinion, the majority asserted:

[U]nder the Due Process Clause of the Fifth Amendment and the notice and jury trial guarantees of the Sixth Amendment, any fact (other than prior conviction) that increases the maximum penalty for a crime must be charged in an indictment, submitted to a jury, and proven beyond a reasonable doubt.<sup>57</sup>

This claim was important, as it suggested that many sentencing guidelines across the country, which relied on factual findings made by a trial judge by a preponderance of evidence, could be subject to future challenge.

---

52. *Id.*

53. *See id.* (surveying state sentencing guidelines).

54. Sentencing Reform Act of 1984, Pub. L. No. 98-473, 98 Stat. 1837, 1987 (1984). Congress created the Federal Sentencing Guidelines through the Sentencing Reform Act of 1984, which "created the United States Sentencing Commission to promulgate the Guidelines." Yang, *supra* note 1, at 1270; *see also* Kate Stith & Steve Y. Koh, *The Politics of Sentencing Reform: The Legislative History of the Federal Sentencing Guidelines*, 28 WAKE FOREST L. REV. 223, 230-60 (1993) (discussing the events leading to eventual passage of the Sentencing Reform Act).

55. 526 U.S. 227 (1999).

56. *Id.* at 229. Describing the question presented in this case, Justice Souter wrote:

This case turns on whether the federal carjacking statute, 18 U.S.C. § 2119, as it was when petitioner was charged, defined three distinct offenses or a single crime with a choice of three maximum penalties, two of them dependent on sentencing factors exempt from the requirements of charge and jury verdict.

57. *Id.* at 243 n.6.

But the Court did not elaborate further. The *Jones* case itself did not provide an obvious procedural vehicle to advance this constitutional argument.<sup>58</sup> Thus, for a short period of time, the footnote provided some glimpse into the Court's thinking without actually altering existing state or federal sentencing laws.

That changed the following year. In *Apprendi v. New Jersey*,<sup>59</sup> the Court found the necessary vehicle to advance the dicta originally expressed in the sixth footnote of *Jones*. The *Apprendi* case considered a New Jersey statute that allowed judges to enhance criminal penalties for crimes deemed “hate crime[s]” when the trial judge found, by a preponderance of evidence, that “[t]he 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.”<sup>60</sup> Charles C. Apprendi Jr. allegedly fired several bullets at the home of a black family that had recently moved into a previously all-white neighborhood.<sup>61</sup> Pursuant to the New Jersey statute, the trial judge held an evidentiary hearing to determine the “purpose” of Mr. Apprendi's actions.<sup>62</sup> At the conclusion of this hearing, the trial judge ruled that the evidence demonstrated, by a preponderance of evidence, that Mr. Apprendi's actions were taken “with a purpose to intimidate” as outlined in the statute, resulting in a sentence of twelve-years imprisonment—significantly above the typical punishment range for this offense of between five- and ten-years imprisonment.<sup>63</sup>

On appeal, Mr. Apprendi argued that the Due Process Clause of the Fourteenth Amendment and the Sixth Amendment guarantee of a jury trial required that any finding used to enhance the statutory maximum sentence be proven to a jury beyond a reasonable doubt.<sup>64</sup> The Court agreed. In ruling the New Jersey statute unconstitutional, the Court noted that the Framers created the right to a jury in the Constitution as a “great bulwark of [our] civil and political liberties.”<sup>65</sup> Historically, the role of the jury was to decide the veracity of all factual allegations against an accused beyond a reasonable

---

58. Specifically, the Court stated later in the footnote that, “[b]ecause our prior cases suggest rather than establish this principle, our concern about the Government's reading of the statute rises only to the level of doubt, not certainty.” *Id.*

59. 530 U.S. 466 (2000).

60. *Id.* at 468–69 (second alteration in original).

61. *Id.* at 469 (noting that Mr. Apprendi made a statement, which he later recanted, saying that even though he did not know the occupants of the house, he did not want them in the neighborhood because of their race).

62. *Id.* at 470 (describing details about this hearing, including the witnesses and evidence presented by both sides).

63. *Id.* at 468–69, 471 (explaining that the typical range is between five and ten years, but also explaining that the hate-crime enhancement could increase this range to between ten and twenty years).

64. *Id.* at 471.

65. *Id.* at 477 (quoting JOSEPH STORY, COMMENTARIES ON THE CONSTITUTION OF THE UNITED STATES 540–41 (4th ed. 1873)) (alteration in original).

doubt.<sup>66</sup> The New Jersey statute effectively transferred this power from the jury, as required by the Constitution, to the trial judge. Ultimately, the Court held that, “[o]ther than the fact of a prior conviction, any fact that increases the penalty for a crime beyond the prescribed statutory maximum must be submitted to a jury, and proved beyond a reasonable doubt.”<sup>67</sup>

After the Court handed down *Apprendi*, there was some uncertainty about the implications of the decision.<sup>68</sup> What did the Court mean when it said the “statutory maximum” for criminal punishment? And how would this understanding of “statutory maximum” apply to jurisdictions that employed sentencing guidelines? For example, in the state of Washington, the legislature established minimum and maximum punishments for each criminal offense. Under state law, offenders found guilty of a class B felony were subject to up to ten-years imprisonment.<sup>69</sup> But judges were not free to issue any sentence between zero and ten years of imprisonment. Instead, judges in Washington were generally bound by sentencing guidelines.<sup>70</sup> Under these guidelines, trial judges had to make a number of factual findings to determine the appropriate punishment for an offender within this zero-to-ten-year statutory range. So for example, even if the statute permitted a punishment of ten years for a class B felony, the sentencing guidelines may require judges to limit their sentence to no more than four years, unless the judge finds the existence of an aggravating factor justifying a departure.<sup>71</sup> What constitutes the statutory maximum punishment under Washington law: the ten-year maximum penalty under the state statute or the hypothetical four-year penalty under the state sentencing guidelines? If a judge made a factual finding to justify a departure from the usual four-year presumptive penalty under the sentencing guidelines, would the fact used to justify this departure need to be submitted to a jury and proven beyond a reasonable doubt in order

---

66. *Id.*

67. *Id.* at 490.

68. Case law immediately after *Apprendi* bolstered this uncertainty. For example, in *Harris v. United States*, the Court at least temporarily suggested that some of these limitations on judicial discretion in sentencing may be permissible after *Apprendi*. 536 U.S. 545, 567–68 (2002). There, the Court held that *Apprendi* did not invalidate mandatory-minimum-sentencing approaches that allowed trial judges to make factual findings by a preponderance of evidence that increased the mandatory minimum sentence within the statutory range but did not alter the maximum sentence under the statute. *Id.* at 568–69. At least temporarily, this led some observers to believe that the long-term implications of *Apprendi* would be relatively minimal.

69. WASH. REV. CODE ANN. § 9A.20.021(1)(b) (West 2000) (current version at WASH. REV. CODE ANN. § 9A.20.021(1)(b) (West, Westlaw through 2015 Wash. Legis. Serv. Ch. 265)) (“No person convicted of a [class B] felony shall be punished by confinement . . . exceeding . . . a term of ten years . . .”).

70. *Blakely v. Washington*, 542 U.S. 296, 299 (2004) (citing the Washington Sentencing Reform Act, which regulated sentences that a trial court judge could give to a criminal defendant under these circumstances).

71. *Id.* (explaining how these sentencing guidelines might apply to a class B felony).

to comply with *Apprendi*? It was not immediately apparent after *Apprendi* whether such a sentencing scheme would violate the Sixth Amendment.

Soon thereafter, the Court tackled this very question in *Blakely v. Washington*. In that case, Ralph H. Blakely pleaded guilty to the kidnapping of his estranged wife.<sup>72</sup> As part of his plea, Blakely confessed to a number of facts that, under the Washington sentencing guidelines, would allow a trial court judge to issue a maximum penalty of fifty-three months in prison—substantially less than the ten-year maximum technically permitted under the requisite Washington statute.<sup>73</sup> But rather than issuing that fifty-three-month prison sentence, the trial court judge found by a preponderance of evidence that Blakely acted with “deliberate cruelty.”<sup>74</sup> Based on this finding, the trial court judge increased Blakely’s punishment to ninety months in prison.<sup>75</sup> That punishment was thirty-seven months above the standard punishment articulated by the sentencing guidelines, but well within the statutorily prescribed maximum sentence of ten years.<sup>76</sup> In reviewing this case, the Court concluded that this arrangement ran afoul of its holding in *Apprendi*. The majority argued that “the ‘statutory maximum’ for *Apprendi* purposes is the maximum sentence a judge may impose *solely on the basis of the facts reflected in the jury verdict or admitted by the defendant.*”<sup>77</sup> Thus, in Washington, the “statutory maximum” for Blakely was fifty-three months in prison. When the trial court judge made a factual finding of deliberate cruelty in order to raise his punishment to ninety months in prison, the trial judge was in direct violation of *Apprendi*. The judge was increasing the maximum allowable sentence under law based on factual findings that were never submitted to a jury or decided beyond a reasonable doubt.

The Court’s decision in *Blakely* had wide-ranging implications. It partially invalidated Washington’s state sentencing guidelines. And the effects of *Blakely* extended beyond Washington. A number of states at the

---

72. *Id.* at 298 (explaining that Blakely kidnapped his wife and detailing the specifics of the offense).

73. *Id.* at 299–300 (“Pursuant to the plea agreement, the State recommended a sentence within the standard range of 49 to 53 months.”).

74. *Id.* at 300 (explaining that the deliberate-cruelty factor was statutorily enumerated in cases of alleged domestic violence).

75. *Id.*

76. WASH. REV. CODE ANN. § 9A.20.021(1)(b) (West 2000) (current version at WASH. REV. CODE ANN. § 9A.20.021(1)(b) (West, Westlaw through 2015 Wash. Legis. Serv. Ch. 265)).

77. *Blakely v. Washington*, 542 U.S. 296, 303 (2004).

time, including Alaska,<sup>78</sup> Arkansas,<sup>79</sup> Florida,<sup>80</sup> Kansas,<sup>81</sup> Michigan,<sup>82</sup> Minnesota,<sup>83</sup> North Carolina,<sup>84</sup> Oregon,<sup>85</sup> and Pennsylvania,<sup>86</sup> used guidelines that were fairly similar to Washington. One estimate by Professor John Pfaff published soon after *Blakely* estimated that the decision would invalidate at least part of the state sentencing-guideline procedures used in thirteen states.<sup>87</sup> Importantly, though, *Blakely* did not bar states from using sentencing guidelines. Instead, it effectively required all states using sentencing guidelines to take one of three paths.<sup>88</sup> First, states could “create a simple, pure or nearly pure ‘charge offense’ or ‘determinate’ sentencing system” where “an indictment would charge a few facts which, taken together, constitute a crime.”<sup>89</sup> Thereafter, any person convicted of that crime would receive the same sentence. Second, legislators could simply abandon all binding sentencing guidelines, opting instead to move back to “indeterminate sentence[s],” or alternatively, voluntary guidelines.<sup>90</sup> After all, *Blakely* only prevented states from using sentencing guidelines that allowed judges to raise the effective maximum sentence for a criminal offense based on any aggravating factors without a jury finding the presence of that aggravating factor beyond a reasonable doubt. Fully voluntary

---

78. ALASKA STAT. § 12.55.155 (2002) (current version at ALASKA STAT. § 12.55.155 (2018) (LEXIS through Alaska Sess. Laws 22)).

79. ARK. CODE ANN. § 16-90-804 (2003 Supp.) (current version at ARK. CODE ANN. § 16-90-804 (LEXIS through 2017 Ark. Acts 423)).

80. FLA. STAT. § 921.0016 (2003) (repealed 2009).

81. KAN. STAT. ANN. § 21-4701 (West 2003) (repealed 2010).

82. MICH. COMP. LAWS ANN. § 769.34 (West 2004), *invalidated in part* by *People v. Lockridge*, 870 N.W.2d 502, 520 (Mich. 2015).

83. MINN. STAT. § 244.10 (2002) (current version at MINN. STAT. ANN. § 244.10 (LEXIS through 2009 Minn. Laws 346)).

84. N.C. GEN. STAT. § 15A-1340.16 (2003) (current version at N.C. GEN. STAT. § 15A-1340.16 (LEXIS through 2017 N.C. Sess. Laws 194)).

85. OR. ADMIN. R. 213-008-0001 (2003), *invalidated* by *State v. Sawatzky*, 96 P.3d 1288 (Or. Ct. App. 2004).

86. 204 PA. CODE § 303.1 (2004) (current version at 204 PA. CODE § 303.1 (LEXIS through 47 Pa. Bull. 5141 Sept. 2, 2017)).

87. Pfaff, *supra* note 7, at 250–51 (“Thus, twelve states besides Washington currently employ sentencing regimes that likely run afoul of *Blakely*.”).

88. *Blakely v. Washington*, 542 U.S. 296, 330 (2004) (Breyer, J., dissenting).

89. *Id.* In his dissent, Justice Breyer worried that, while a system “assures uniformity,” it does so “at intolerable costs” because it would result in individuals being given identical sentences for behavior that happened under vastly different circumstances. *Id.* at 330–31. Justice Breyer also worried that such a statutorily fixed mandatory sentencing regime would shift “tremendous power to prosecutors to manipulate sentences through their choice of charges.” *Id.* at 331. This risk was heightened, according to Breyer, because of the reality that most cases are resolved via plea bargaining. *Id.*

90. *Id.* at 332. Justice Breyer went on to criticize these systems, claiming that indeterminate sentencing regimes almost always lead to disparities based on everything from race to what a judge may have eaten for breakfast that day. *Id.*

guidelines do not violate *Blakely*.<sup>91</sup> Or third, states could continue to employ presumptive sentencing guidelines, so long as they procedurally complied with *Blakely*.<sup>92</sup> Such attempts to make existing presumptive sentencing guidelines compliant with *Blakely* would require trial courts to submit to a jury any aggravating factor that could cause a sentence to exceed the presumptive-guideline range, other than factors like prior criminal convictions.<sup>93</sup>

But as discussed more in the next subpart, dissenters in *Blakely* and scholars alike issued dire predictions about how *Blakely*-compliant guidelines would affect sentence lengths and sentence disparities.

### B. Predictions About *Blakely*-Compliant State Sentencing Guidelines

In the wake of *Blakely*, many worried that the decision would undo decades of progress in criminal justice reform. First, in the immediate aftermath of *Blakely*, some critics worried that at least some states would simply move away from presumptive sentencing guidelines, opting instead for voluntary guidelines, or even fully discretionary sentencing within statutorily prescribed ranges. This, critics argued, would ultimately lead to more disparity in criminal sentences.<sup>94</sup>

---

91. Under such an approach, a judge would have complete discretion to impose any sentence up to the maximum prescribed by the statute. This may allow for sentences to be highly individualized to the relative blameworthiness of each offender. And some guidelines, even if they are voluntary, may still be preferable to no guidelines at all. But such a voluntary system comes with obvious drawbacks—namely, that without binding guidance, judges may ultimately give in to bias, resulting in significant disparities in sentences based on race, gender, class, or other impermissible factors. To be clear, not all disparities between sentences are the result of any bias on the part of the judge. Two judges may have legally defensible but divergent sentencing philosophies. This may lead to different judges issuing sentences that differ from one another on the basis of judicial philosophy rather than considerations of any impermissible factor. Empirically distinguishing between the two is a methodological challenge.

92. See *Blakely*, 542 U.S. at 333 (Breyer, J., dissenting) (discussing how states may retain presumptive sentencing guidelines if they are compliant with *Apprendi* and by extension *Blakely*).

93. In order to be compliant with *Blakely* and the Sixth Amendment, the jury must then find the presence of the aggravating factor beyond a reasonable doubt. *Id.* at 301 (majority opinion). Such a *Blakely*-compliant guideline system could still permit trial judges to reduce sentence lengths based on the finding of mitigating factors. Such mitigating factors need not be found beyond a reasonable doubt.

94. Justice O'Connor and other dissenters in the case made just such an argument. Before Washington introduced its sentencing guidelines in 1981, Justice O'Connor recalled how "[s]entencing judges, in conjunction with parole boards, had virtually unfettered discretion to sentence defendants to prison terms falling anywhere within the statutory range, including probation—*i.e.*, no jail sentence at all." *Id.* at 315 (O'Connor, J., dissenting). This "system of unguided discretion" contributed to "severe disparities in sentences received and served by defendants committing the same offense and having similar criminal histories." *Id.* And frequently, these disparities tracked constitutionally suspect variables like race. *Id.* The Washington guidelines were a direct response to this well-documented problem, not an attempt by the legislature to circumvent the procedural protections guaranteed by the Sixth Amendment of the Constitution. And as Justice O'Connor observed, the Washington guidelines appeared to achieve their intended goal. The state experienced a substantial reduction in apparent signs of sentencing disparities. *Id.* at 317–



Second, and relatedly, critics argued that state attempts to maintain sentencing guidelines while complying with *Blakely* would prove unnecessarily burdensome and taxing to implement.<sup>95</sup> While the majority in *Blakely* pointed out that its decision did not prevent states from using mandatory or presumptive sentencing guidelines, Justice O'Connor noted that the decision would nonetheless "exact[] a substantial constitutional tax" on states by forcing them to conduct "full-blown jury trial[s] during the penalty phase proceeding[s]" to determine an offender's sentence.<sup>96</sup> And according to Justice O'Connor and the dissenters, "simple economics dictate" that at least some states would not be able to bear all of these attendant costs.<sup>97</sup> Again, scholars echoed these critics.<sup>98</sup>

And third, at least some critics worried that attempts to make sentencing guidelines *Blakely*-compliant would ultimately contribute to a more punitive justice system. Perhaps most notably, Professor Steven L. Chanenson predicted that this approach to sentencing guidelines may result in an overall increase in the severity of criminal sentences. If a jury found the presence of an aggravating factor beyond a reasonable doubt, Chanenson worried that "it [would be] asking a great deal of any judge, particularly an elected judge, to exercise her discretion and deny that departure."<sup>99</sup> While many judges regularly denied upward departures in sentences under purely voluntary guidelines, Chanenson worried that *Blakely*-ized guidelines would ultimately pressure judges to sentence more severely. Relatedly, Chanenson expressed concern that *Blakely*-ized guidelines may contribute to legislative efforts to increase the overall severity of the criminal justice system.<sup>100</sup> No sentencing guidelines can possibly capture all factors that speak to the moral blameworthiness of an individual offender. So, to the extent that *Blakely* requirements limit the ability of judges to depart upward in sentencing for particularly egregious criminal offenders, Chanenson worried that *Blakely*-

---

18. Thus, some understandably worried that the decision in *Blakely* may contribute to greater disparities in sentencing as it rolled back the ability of states to require judges to make factual findings that altered sentence lengths, and scholars at the time shared this concern. See Nancy J. King & Susan R. Klein, *Beyond Blakely*, 16 FED. SENT'G REP. 316, 318 (2004) (arguing that *Blakely* hinders efforts to reduce sentencing disparity). It is also worth mentioning that Justice O'Connor's description of the disparities that existed before the imposition of sentencing guidelines in Washington fits in large part Alabama's prestandards and, to a significant extent, Alabama's voluntary standards paradigms.

95. Justice O'Connor in particular argued that this so-called *Blakely*-ization of sentencing guidelines brings with it "substantial and real" costs. *Blakely*, 542 U.S. at 318 (O'Connor, J., dissenting).

96. *Id.* at 318–19.

97. *Id.* at 320.

98. For example, one scholar argued that post-*Blakely*, "too much judicial discretion . . . will lead to increased (and perhaps invidious) disparity." Chanenson, *supra* note 40, at 431.

99. *Id.* at 425.

100. *Id.*

ized guidelines may result in state legislators instead “shift[ing] the entire system up a notch or two in severity” in response.<sup>101</sup>

To be clear, *Blakely* does not prevent states from utilizing sentencing guidelines. But there is at least some evidence to suggest, consistent with the predictions made by many critics, that some states moved away from presumptive sentencing guidelines after *Blakely*, opting instead for voluntary guidelines or no guidelines. Nevertheless, as explained in more depth in the next subpart, little existing research has explored the comparative advantages of voluntary and presumptive sentencing guidelines at the state level.

### C. Existing Research

Several studies have explored the effects of federal sentencing guidelines. Specifically, research has found that the introduction of federal sentencing guidelines contributed to reductions in interjudge sentencing disparities and racial disparities. Similarly, some research has found that *United States v. Booker*, which effectively invalidated the mandatory nature of the Federal Sentencing Guidelines for the same reasons as *Blakely*,<sup>102</sup> resulted in a subsequent increase in racial and interjudge disparities in the federal system. At least one study, published soon after the *Blakely* decision, has considered the effectiveness of voluntary sentencing guidelines relative to mandatory sentencing guidelines at the state level. But the overall literature on state sentencing guidelines leaves considerable room for additional research. Below we provide an admittedly brief summary of some of these important studies.

First, some researchers have studied how the introduction of the Federal Sentencing Guidelines influenced judicial behavior.<sup>103</sup> These studies have commonly found that the introduction of the Federal Sentencing Guidelines was associated with statistically significant reductions in interjudge disparities,<sup>104</sup> even if these effects were felt unevenly across districts.<sup>105</sup> And

---

101. *Id.* Chanenson also worried that other attempts to *Blakely*-ize guidelines might more technically comply with *Blakely*, but ultimately “unregulated discretion” would only operate to “lengthen sentences.” *Id.* at 418–19.

102. See *United States v. Booker*, 543 U.S. 220, 235, 244 (2005) (holding that the Federal Sentencing Guidelines impermissibly gave trial judges the authority to make factual findings by a standard lower than beyond a reasonable doubt and use these determinations to raise the overall maximum sentence length in criminal cases in violation of the Sixth Amendment).

103. Paul J. Hofer et al., *The Effect of the Federal Sentencing Guidelines on Inter-Judge Sentencing Disparity*, 90 J. CRIM. L. & CRIMINOLOGY 239, 241 (1999).

104. For example, in 1999, Hofer, Blackwell, and Ruback found that the Federal Sentencing Guidelines “significantly reduced overall inter-judge disparity in sentences imposed, much as the parole guidelines reduced disparity in time actually served prior to implementation of the sentencing guidelines.” *Id.*

105. *Id.* (“Some types of cases show no improvement, or show improvement in some cities but not in others. Further, there is evidence that some regional sentencing differences have increased under the guidelines, particularly in drug trafficking cases.”). Additionally, scholars James M. Anderson, Jeffrey R. Kling, and Kate Stith published a study analyzing the effect of the Federal

some prior research conducted both before and after *Booker* and *Blakely* has found that factors like the judge's race, gender, or political affiliation exert some influence on the judge's sentencing behavior and exercise of discretion.<sup>106</sup>

Second, a handful of recent studies have examined the effect of the *Booker* decision invalidating the mandatory nature of the Federal Sentencing Guidelines.<sup>107</sup> These studies have generally found that by moving the federal criminal system from binding to voluntary sentencing guidelines, the *Booker* decision contributed to an increase in interjudge and racial disparities in criminal sentencing.<sup>108</sup> For example, in 2014, Professor Crystal S. Yang

---

Sentencing Guidelines on sentence disparities. See generally James M. Anderson et al., *Measuring Interjudge Sentencing Disparity: Before and After the Federal Sentencing Guidelines*, 42 J.L. & ECON. 271 (1999). They found that, after the passage of the Federal Sentencing Guidelines, the average interjudge disparity dropped from around 4.9 months to around 3.9 months. *Id.* at 294.

106. See, e.g., Joshua B. Fischman & Max M. Schanzenbach, *Do Standards of Review Matter? The Case of Federal Criminal Sentencing*, 40 J. LEGAL STUD. 405, 406 (2011) (finding judges appointed by Democrats are more lenient than those appointed by Republicans); Max M. Schanzenbach, *Racial and Sex Disparities in Prison Sentences: The Effect of District-Level Judicial Demographics*, 34 J. LEGAL STUD. 57, 72–73 (2005) (finding racial minority and female judges differ from others in sentencing); Max M. Schanzenbach & Emerson H. Tiller, *Reviewing the Sentencing Guidelines: Judicial Politics, Empirical Evidence, and Reform*, 75 U. CHI. L. REV. 715, 734 (2008) (finding that judges appointed by Republicans give longer sentences than those appointed by Democrats); Max M. Schanzenbach & Emerson H. Tiller, *Strategic Judging Under the U.S. Sentencing Guidelines: Positive Political Theory and Evidence*, 23 J.L. ECON. & ORG. 24, 52–53 (2007) (finding that judges appointed by Democrats issue lower sentences for serious crimes than those appointed by Republicans); Susan Welch et al., *Do Black Judges Make a Difference?*, 32 AM. J. POL. SCI. 126, 134 (1988) (concluding that race has some effect on sentences issued by judges).

107. For example, in 2010, Professor Ryan W. Scott found that, by making the federal guidelines advisory in *Booker*, setting highly deferential standards of review in *Gall v. United States*, 552 U.S. 38 (2007), and authorizing judges to reject the use of the sentencing guidelines based on policy preferences in *Kimbrough v. United States*, 552 U.S. 85 (2007), the U.S. Supreme Court likely contributed to significant increases in interjudge disparities in the federal criminal system. Ryan W. Scott, *Inter-Judge Sentencing Disparity After Booker: A First Look*, 63 STAN. L. REV. 1, 3, 30–41 (2010). To prove this, Professor Scott relied on data from the District of Massachusetts, which at the time was the only district court to make this sort of information publicly available. *Id.* at 21–24 (explaining the source of the data and further explaining that the data set was made possible through using “a method, pioneered by Max Schanzenbach and Emerson Tiller, that matches publicly available docket information with corresponding information in the Commission’s case records”). His results found “a clear increase in inter-judge sentencing disparity, both in sentence length and in guideline sentencing patterns.” *Id.* at 30. He also found this effect “doubled in strength” after *Kimbrough* and *Gall*. *Id.*

108. For example, in 2012, the U.S. Sentencing Commission released a report on the effects of *Booker* on federal sentencing. U.S. SENTENCING COMM’N, REPORT ON THE CONTINUING IMPACT OF UNITED STATES V. BOOKER ON FEDERAL SENTENCING (2012), <https://www.ussc.gov/research/congressional-reports/2012-report-congress-continuing-impact-united-states-v-booker-federal-sentencing> [<https://perma.cc/23DQ-CDY7>]. This extensive, multipart report considered a range of independent variables and conducted multivariate regressions using all cases up to 2011 to determine whether *Booker* and its progeny impacted an offender’s total sentence length. *Id.* pt. E, at 7–8. It ultimately concluded that “sentencing outcomes increasingly depend upon the district in which the defendant is sentenced” and that “[d]emographic factors (such as race, gender, and citizenship) [have been] associated with sentence length at higher rates in the *Gall* period than in

carried out a prominent study on the effects of *Booker*'s limitation on the use of the Federal Sentencing Guidelines.<sup>109</sup> By analyzing data from between 2000 and 2009,<sup>110</sup> she found that a "defendant who is randomly assigned to a one-standard-deviation 'harsher' judge in the district court received a 2.8-month longer prison sentence compared to the average judge before *Booker*, but received a 5.9-month longer sentence [after *Booker*] . . . , a doubling of interjudge disparities."<sup>111</sup> Relatedly, some studies have debated how to empirically differentiate the effects of judicial discretion from the effects of offender behavior and prosecutorial discretion. Because of this, some scholars have argued that the best available empirical evidence does not necessarily support the claim that *Booker* increased racial disparities in sentencing in the federal criminal system.<sup>112</sup> However, this claim remains subject to scholarly debate.<sup>113</sup>

Finally, a relatively smaller body of research has explored the comparative effectiveness of different sentencing-guideline models at the state level. One of the best studies to date on this topic comes from a 2006 article by Professor John Pfaff, which relied on data from the National Corrections Reporting System to estimate the comparative usefulness of voluntary and presumptive sentencing guidelines across a number of states.<sup>114</sup> Professor Pfaff found that a state's use of voluntary guidelines

---

previous periods." *Id.* pt. A, at 89, 108. Also in 2012, Professors Susan B. Long and David Burnham published a study using data from over 370,000 criminal cases compiled by the Transactional Records Access Clearinghouse. Susan B. Long & David Burnham, *TRAC Report: Examining Current Federal Sentencing Practices: A National Study of Differences Among Judges*, 25 FED. SENT'G REP. 6, 6 (2012). They found statistically significant disparities in the sentences given by different judges to similarly situated defendants. *Id.* at 15.

109. See generally Yang, *supra* note 1, at 1268 (utilizing data from 400,000 criminal defendants linked to sentencing judges to analyze interjudge disparities after *Booker*).

110. Specifically, Professor Yang's study was "the first in over twenty-five years to match sentencing data to judge identifiers in all ninety-four district courts, allowing for a comprehensive look at interjudge sentencing disparities after *Booker*." *Id.* at 1294.

111. *Id.* at 1275. And she found that the same proved true for more lenient judges, as "a defendant randomly assigned to a one-standard-deviation more 'lenient' judge faced a 4.7% chance of receiving a below-range departure before *Booker*, but over a 6.9% chance after *Kimbrough/Gall*." *Id.*

112. 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, 2 (2013). Using "rigorous regression discontinuity-style design," Professors Starr and Rehavi found little evidence to suggest that *Booker* caused any significant increase in sentencing disparities. *Id.* Instead, they have attributed much of the existing sentencing disparities in the federal system to discretionary decisions made by prosecutors, not judges. M. Marit Rehavi & Sonja B. Starr, *Racial Disparity in Federal Criminal Sentences*, 122 J. POL. ECON. 1320, 1320 (2014).

113. Specifically, it is worth noting that the Federal Sentencing Commission has issued a response to Professors Starr and Rehavi's work in which the Commission expressed disagreement with many of their methodological choices. Glenn R. Schmitt et al., *Why Judges Matter at Sentencing: A Reply to Starr and Rehavi*, 123 YALE L.J. ONLINE 251, 251–52 (2013).

114. Pfaff, *supra* note 7, at 235, 256. During this time period, Pfaff identified some states that fell into multiple different categories: (1) some states either enacted during this time period or employed throughout this time period voluntary sentencing guidelines; (2) some states either

resulted in a reduction of variation in sentence length by as much as 35% for violent crimes and 21% for property crimes relative to jurisdictions without sentencing guidelines.<sup>115</sup> By comparison, presumptive or mandatory sentencing guidelines resulted in reductions in the variation of sentence lengths for similarly situated defendants of around 57% for violent crimes and 54% for property crimes relative to jurisdictions without sentencing guidelines.<sup>116</sup> Based on this finding, Professor Pfaff concluded that voluntary sentencing guidelines are able to reduce sentence variation almost, but not quite as much as, binding sentencing guidelines.<sup>117</sup> Thus, he hypothesizes that the effects of *Blakely* on sentence variations in some states may be less than initially anticipated—provided that jurisdictions replaced binding sentencing guidelines with voluntary guidelines.<sup>118</sup> While we address a similar research question to Professor Pfaff, as explained in more depth in the next Part, we use a different methodological approach that allows us to build on his findings and add to the existing literature.<sup>119</sup>

Admittedly, this only scratches the surface of the many important studies on sentencing guidelines. But overall, the existing literature leads to

---

enacted during this time period or employed throughout this time period binding sentencing guidelines; and (3) some states used no sentencing guidelines during this time period. *See id.* at 250–54 (observing the diverse nature of state sentencing guidelines).

115. *Id.* at 235.

116. *Id.*

117. *Id.* at 268.

118. *Id.* at 283–85.

119. Specifically, we believe that our study’s methodology differs in three significant ways. These differences give us additional insight into this empirical question. First, rather than attempting to compare the experiences of different states using NCRS data, we instead focus on the experience of one single state (Alabama) that has experimented with different forms of sentencing guidelines over the last two decades. The Alabama data allow us to compare the behavior of judges within a single jurisdiction over time in response to new forms of external regulation. Second, and relatedly, our methodological approach also allows us to track the behavior of *individual judges* over time in a single jurisdiction in response to the introduction of new sentencing guidelines. We do this in a couple of ways: we look at how the same judges treat similarly situated defendants under different sentencing-guideline regimes, and we look at how the same judges treat defendants covered by these sentencing guidelines with defendants that are not covered by these sentencing guidelines. We think this methodology allows us to make more confident predictions about how judges respond to the introduction of different sentencing guidelines. This is consistent with many of the studies involving the behavior of federal judges after *Booker*. And third, while Professor Pfaff’s study explored whether voluntary sentencing guidelines could help states make up some of the ground they lost after *Blakely* overturned some types of binding sentencing guidelines on Sixth Amendment grounds, our study tackles this issue in a somewhat different way. Remember that at the time of the *Blakely* decision, Alabama did not employ sentencing guidelines. In the years that followed, Alabama began experimenting with a couple forms of sentencing guidelines that complied with the requirements of *Blakely*—first voluntary guidelines, then presumptive guidelines. By exploring how these different guideline approaches influenced the sentences handed down by the largely same group of judges over time, our study allows us to examine the efficacy of different constitutionally permissible approaches to sentencing guidelines after *Blakely*. It helps us understand whether states that use presumptive sentencing guidelines after *Blakely* can still effectively control judicial discretion and sentencing disparities. Our study ultimately builds and extends on Professor Pfaff’s excellent and important work.

a couple of hypotheses. The existing literature generally suggests that mandatory or presumptive sentencing guidelines at the state and federal levels have been reasonably effective at altering judicial behavior. These types of binding sentencing guidelines are associated with reductions in disparities in sentences, both between judges and based on race. But there remains some debate about the relative effectiveness of voluntary guidelines as compared to presumptive guidelines. And while at least one study suggests that voluntary sentencing guidelines may be able to provide some of the benefits of mandatory guidelines, there is need for more research on the comparative usefulness of different types of state sentencing guidelines, particularly post-*Blakely*. As discussed in more depth in the next Part, this study helps fill some of these gaps in the existing literature.

## II. Methodology

This Article seeks to examine the effects of voluntary and presumptive sentencing guidelines in Alabama on sentence length, sentence disparities, and racial disparities. To answer the empirical questions at hand, this Article relies on a comprehensive and nonpublic data set of all criminal cases from Alabama between 2002 and 2015. As the subparts that follow describe in more detail, Alabama provides a rare opportunity to explore the effects of both voluntary and presumptive sentencing guidelines because of its unique legislative history. Subpart A provides background on Alabama's history with sentencing guidelines between 2002 and 2015. Subpart B walks through the data set used in this study. And subpart C presents three different models that we employ to evaluate the effects of these different sentencing-guideline structures in Alabama.

### A. *Background on Alabama's Sentencing Guidelines*

We chose to focus on Alabama in this study because of its unique history with sentencing guidelines over the last two decades. As best we can tell, Alabama is one of the only states in recent history to experiment with both voluntary and presumptive sentencing guidelines over a relatively short period of time. Until 2006, Alabama employed no sentencing guidelines.<sup>120</sup> Between 1970 and 2000, Alabama saw incarceration increase by 326%, and the state's incarceration rates per capita ranked well above the national average, leading to concern among many policymakers.<sup>121</sup> State leaders also recognized that "unwarranted sentencing disparities" led to concerns about

---

120. See ALA. SENTENCING COMM'N, 2007 JUDGES' SENTENCING REFERENCE MANUAL 75 (2007) [hereinafter 2007 MANUAL], [http://sentencingcommission.alacourt.gov/publications/judges%20reference%20manual\\_july2007.pdf](http://sentencingcommission.alacourt.gov/publications/judges%20reference%20manual_july2007.pdf) [https://perma.cc/729Y-YE7H] (explaining that Alabama first instituted sentencing guidelines in 2006).

121. *Id.* ("Alabama had twice as many property crimes admissions per 100 arrests between 1983 and 1992 as the national average. Drug offenders represent the largest percentage of offenders entering Alabama prisons.").

fairness in sentencing procedures.<sup>122</sup> So soon thereafter, the state legislature passed the Sentencing Reform Act of 2003, which authorized the Alabama Sentencing Commission to create sentencing guidelines, which the legislature later approved in 2006.<sup>123</sup> Initially, the Sentencing Commission only recommended the implementation of voluntary sentencing guidelines, which were designed to maintain “judicial discretion and sufficient flexibility to permit individualized sentencing as warranted by mitigating and aggravating factors.”<sup>124</sup>

Unlike many states that utilize grid systems, Alabama’s guidelines take the form of a worksheet. The state utilizes three different types of worksheets: (1) *personal worksheets*, which cover crimes like assault, manslaughter, murder, rape, and robbery; (2) *drug worksheets*, which cover felony DUI, the manufacturing of a controlled substance, possession of a controlled substance, and sale, distribution, or intent to distribute a controlled substance; and (3) *property worksheets*, which cover burglary, theft of property, vehicle theft, forgery, and other similar offenses.<sup>125</sup> Beginning in 2006, Alabama required trial court judges to “indicate on the record that the worksheet and applicable sentencing standards have been reviewed and considered.”<sup>126</sup> Some offenses, like sexual offenses committed against children under the age of twelve, contraband manufacturing, obstruction of justice, and some additional domestic- and sexual-abuse offenses, are considered *nonworksheet offenses* as they have not been subject to any voluntary or presumptive worksheet over this time period.<sup>127</sup>

The worksheets use the class and severity of the crime, the number of prior convictions, incarceration, probation, use of a weapon, and injury to the

---

122. *Id.* at 75–76 (“[T]he Alabama legislature has created the Alabama Sentencing Commission to recommend changes in Alabama’s criminal justice system. Such recommendations must, among other things, secure public safety, provide certainty and fairness in sentencing, avoid unwarranted sentencing disparities, and prevent prison overcrowding and premature release of prisoners.”).

123. *Id.* at 75. The Commission further elaborated that “[t]he recommendations or ‘standards’ as they are called are voluntary, non-appealable, historically based, time imposed, sentencing recommendations developed for 26 felony offenses, representing 87% of all felony convictions and sentences imposed in Alabama over an approximate five-year period from October 1, 1998 through May 31, 2003.” *Id.*

124. *Id.* at 76 (quoting ALA. CODE § 12-25-2 (2007)). The *Blakely* decision influenced this choice by the Alabama Sentencing Commission—particularly in light of the fact that the Commission was instructed to maintain judicial discretion. *See supra* note 91.

125. *See* 2016 MANUAL, *supra* note 23, at 21–22 (showing the different worksheets available after the passage of the presumptive sentencing guidelines); *see also* ALA. SENTENCING COMM’N, 2005 REPORT app. at 1 (2004) [hereinafter 2005 REPORT], <http://sentencingcommission.alacourt.gov/media/1043/2005-annual-report.pdf> [<https://perma.cc/K9V9-M4PL>] (showing the initial voluntary sentencing guidelines worksheet categories that originally went into effect in 2006).

126. 2007 MANUAL, *supra* note 120, at 76 (quoting ALA. CODE § 12-25-35 (2007)).

127. *See* 2005 REPORT, *supra* note 125, app. at 1 (listing the offenses covered by the voluntary worksheets); 2016 MANUAL, *supra* note 23, at 22 (same). Nonworksheet offenses are addressed either by discretionary judicial sentencing or by legislatively imposed mandatory sentencing.

victim to calculate a score for each offender.<sup>128</sup> Upon calculating a score, the voluntary worksheets give the judge an upper and lower boundary for total and imposed sentence length.<sup>129</sup> While the Commission encouraged judges to stay within these boundaries for total and imposed sentences, these initial worksheets were merely voluntary; judges could follow the worksheet's guidelines, or they could choose to depart from the recommended sentence lengths without facing significant scrutiny if they felt that the recommendations failed to match the offender's culpability.<sup>130</sup>

Sentencing in Alabama proceeded in this manner until October 2013. At that point:

[T]he Alabama Legislature changed the Standards for non-violent offenses [i.e. property and non-violent drug offenses] . . . from voluntary to presumptive recommendations and directed the Alabama Sentencing Commission to make modifications as necessary to effect this change, including defining aggravating and mitigating circumstances that are required for sentencing departures from presumptive recommendations.<sup>131</sup>

Again, in making these changes, Alabama policymakers stated that the goal was to reduce “unwarranted disparity and prison overcrowding” so as to reserve “scarce prison resources for the most dangerous and violent offenders.”<sup>132</sup> But even after the legal changes in 2013, adherence to the guidelines remained voluntary for violent offenses, including many personal offenses, burglary offenses, and violent drug offenses.<sup>133</sup> And so-called nonworksheet offenses were not bound by any guidelines.<sup>134</sup>

Figure 1 below provides a graphical summary of the history of sentencing-guideline experimentation by the state of Alabama between 2002 and 2015, the time period for our study.

---

128. 2007 MANUAL, *supra* note 120, at 76–80 (showing a detailed breakdown of how judges ought to process this new worksheet soon after its release).

129. *Id.* at 78 (describing the details on compliance and the types of sentences that judges could award under these new standards).

130. *Id.* at 75–80 (emphasizing multiple times that the guidelines were merely voluntary or advisory).

131. 2016 MANUAL, *supra* note 23, at 15 (further elaborating that the offenses involved in the new presumptive sentencing guidelines were those covered by ALA. CODE § 12-25-32).

132. *Id.*

133. *Id.* at 16.

134. *Id.* at 24.



*Figure 1: Alabama Sentencing Guidelines Over Time*

January 2002 to October 2006	October 2006 to October 2013	October 2013 to December 2015
<u>No sentencing guidelines</u> for all crimes	<u>Voluntary guidelines</u> (1) personal crimes (2) burglary (3) nonviolent drug crimes (4) nonviolent property crimes  <u>No sentencing guidelines</u> for other crimes	<u>Voluntary guidelines</u> (1) personal crimes (2) burglary  <u>Presumptive guidelines</u> (1) nonviolent drug crimes (2) nonviolent property crimes  <u>No sentencing guidelines</u> for other crimes

This unique series of legal events in a single state makes this a useful opportunity to explore the effect of voluntary and presumptive sentencing guidelines on judicial behavior for a few reasons. For one thing, it means that we can examine the effect of both voluntary and presumptive guidelines on sentencing outcomes within a single legal community. While the Alabama Sentencing Commission and the legislature established and then expanded the use of sentencing guidelines between 2002 and 2015, we have been unable to identify any other significant changes to the legal landscape during this time period that would significantly affect judicial sentencing behavior.<sup>135</sup> And since we are not comparing one group of states to another

---

135. To be clear, numerous factors other than the presence of sentencing guidelines affect criminal sentences. For example, the actions of prosecutors may have a substantial effect on criminal sentences. Our claim here is narrower. Even if there are other factors that may affect sentence length, we have been unable to identify any factors that *changed* between 2002 and 2015 other than the implementation of these sentencing guidelines. This allows us to be more confident that any subsequent changes we observe in sentences handed down by judges are the result of these changes in sentencing guidelines and not other variables that merely correlated with the introduction of these sentencing guidelines. Additionally, some may point out that during a fourteen-year period, some judges resigned, retired, were removed, or died. While this is true, we do not believe this affects our analysis. For one thing, we are able to track individual judges to see not just how overall sentence length changes during this time period but also how individual judges who served during this *entire time period* changed their behavior. Additionally, to the extent that the judicial election process may influence sentencing behavior, we have no reason to believe that these procedures exerted a substantially different influence on judicial behavior at the start of our time period relative to the end of our time period. Thus, even if these factors have some influence on overall sentence lengths or overall disparities, they likely exerted this same effect throughout the fourteen-year period—meaning that the changes we observe are more likely the result of the dramatic introduction of sentencing guidelines rather than background variables that remain constant throughout this period.

(as some prior studies have done),<sup>136</sup> we feel somewhat more confident that we can attribute changes in judicial behavior to changes in sentencing procedures, rather than other, difficult-to-measure variables. For example, it is theoretically plausible that plea-bargaining procedures, prosecutorial norms, or the process by which judges are elected or appointed to the bench may all influence sentence lengths and disparities. In studies that compare one state to another state, controlling for these alternative explanatory variables is challenging. Without evidence of any other significant changes in the legal landscape in Alabama that may influence sentencing behavior, we believe that the introduction of sentencing guidelines is the most likely causal contributor to sudden, subsequent changes in judicial sentencing behavior that we illustrate *infra* Part III.

Additionally, because these changes happened over a relatively short historical time period in a single state, the community of judges in Alabama remained relatively unchanged. This allows us to examine how each legal change in the state influenced the judicial behavior, particularly as our data set allows us to track 355 individual judges over this fourteen-year period, as described in the next subpart. We are also able to use the class of offenses that have never been subject to sentencing guidelines as a control, providing us a baseline for comparison.

### B. *Data Set*

This Article draws on a data set of 221,934 criminal sentences handed down by 355 different judges in Alabama between 2002 and 2015 to explore the effects of voluntary and presumptive sentencing guidelines post-*Blakely*. The Alabama Sentencing Commission provided the data for this study.<sup>137</sup> Our data set contains virtually every criminal sentence between 2002 and 2015. The main outcomes of interest in our study are the total sentence length<sup>138</sup>

---

136. See generally Pfaff, *supra* note 7 (comparing states that utilize voluntary guidelines, presumptive guidelines, and no guidelines in order to judge the effect of *Blakely* on future sentencing outcomes).

137. The Commission released this information under strict requirements limiting public disclosure and public identification of cases, defendants, judges, prosecutors, or geographical information. The Commission also released these data to us with the understanding that it retained all rights to the database.

138. Admittedly, the term “total sentence length” may be up for some debate. We define this term to be the total length of incarceration assigned by the trial court judge in a given case, regardless of whether these sentences are served in prison or jail. So, as a hypothetical scenario, a judge could sentence a felon to four years, but only impose two of those four years, so a convicted individual serves two years in prison or jail with a period of probation (i.e., a “split sentence” in Alabama terminology). If the felon, for instance, violates the terms of probation set by the judge, the original total sentence may become binding, requiring the felon to serve the full four-year term. But for our purposes, we define this sentence handed down by the judge as a two-year term.

handed down by the trial court judge and how that sentence length compared to similarly situated defendants.<sup>139</sup>

Of course, a number of different variables may influence the length of a criminal sentence handed down by a judge. For example, the seriousness of the offense or circumstances, the personality of a trial court judge, the offender's previous criminal history, any recommendation by the state prosecutor, and a host of other factors may independently influence the length of a criminal sentence. The goal of this study is to assess the effect of voluntary and presumptive sentencing guidelines on judicial behavior, controlling for any other possible explanations. To address this concern, our models control for as many potentially explanatory variables as possible. The data set collected from the Alabama Sentencing Commission allows us to control for many of these important, alternative explanations. Our data set includes many case-specific characteristics, including the seriousness of the conviction offense, seriousness of the indictment offense,<sup>140</sup> whether or not a defendant agreed to a plea bargain, the number of counts in an indictment, the number of prior offenses, whether or not the ruling was a split sentence involving a period of probation, if a drug or mental health court was used, if the ruling included a requirement to attend a drug treatment or counseling program, and if the conviction included drug activity near schools or housing projects. Our data set also allows us to control for demographic characteristics of the defendant, such as race and gender. Additionally, we are able to include fixed effects for judges, circuits, counties, the most serious offense at indictment, and the most serious offense at conviction. And we have historical records of past sentences handed down by each judge. Thus, we know whether a judge has previously issued particularly harsh or lenient sentences relative to his or her peers. This allows us to better understand whether the sentencing guidelines have different effects on different types of judges, based in part upon that judge's preexisting tendency towards harsh or lenient sentences.

---

139. Our data set does not document events that happen to an offender after a judge hands down a criminal sentence. For example, we do not have data on whether an inmate receives parole at some point while serving a criminal sentence. Our data set merely includes the sentences handed down by the trial judge and information on the judge, defendant, criminal offense, and any applicable sentencing guidelines. Ultimately, we do not believe that our lack of data on subsequent decisions by the Alabama Board of Pardons and Paroles (ABPP) limits the findings of this study. This study aims to understand how external legal regulations affect the behavior of trial judges in handing down criminal sentences. Thus, we are primarily concerned about the total sentence length handed down by the trial court judge and how that sentence length compared to similarly situated defendants.

140. Seriousness of the indictment and conviction offenses is calculated by a scoring technique developed and used by the Alabama Sentencing Commission and is used on the worksheet score calculations. For more information, see ALA. SENTENCING COMM'N, <http://sentencingcommission.alacourt.gov> [<https://perma.cc/PDR3-3ZX4>] (navigate to "publications" and select the appropriate worksheet for review). See, e.g., 2016 MANUAL *supra* note 23, at 39 (assigning numerical values to certain offenses).

We recognize that this does not exhaust all possible explanatory variables. Nevertheless, we have found that this data set is roughly consistent with other prior studies.<sup>141</sup>

### C. Models

This study presents three separate models, each designed to analyze a somewhat different empirical question. First, we present below a model designed to evaluate the effect of voluntary and presumptive sentencing guidelines on the length of sentences. Second, we present below a model that explores whether the introduction of voluntary and presumptive sentencing guidelines affected the level of racial disparities in sentences. And third, we offer a model to examine the effect of these guidelines on interjudge disparities. We discuss each model in turn.

1. *Measuring the Effect on Sentence Length.*—To estimate the effect of the introduction of different types of sentencing guidelines on sentence lengths, we employ the following model:

$$O_{it} = a_1 \text{Group}V_{it} + a_2 \text{Post}V_{it} + \alpha \text{Group}V_{it} * \text{Post}V_{it} + b_1 \text{Group}P_{it} + b_2 \text{Post}P_{it} + \beta \text{Group}P_{it} * \text{Post}P_{it} + X_{it} + \varepsilon_{it} \quad (1)$$

which, for each policy change, includes postpolicy time dummy variables—*PostV* and *PostP*—that flag the postpolicy periods of time for each respective change, group dummy variables—*GroupV* and *GroupP*—and the corresponding interaction terms to get our difference-in-differences estimators of interest,  $\alpha$  and  $\beta$ .

The variable *GroupV* represents a dummy variable for any conviction that is subject to voluntary sentencing guidelines. This includes convictions in the personal and burglary class of offenses and drug or property class of offenses.<sup>142</sup> The variable *GroupP* represents a dummy for just those nonviolent crimes that ultimately are deemed to fall under presumptive sentencing guidelines. Additionally, a host of controls, *X*—which includes

141. See, e.g., Pfaff, *supra* note 7, at 255–68 (describing data set and methodology); Yang, *supra* note 1, at 1294–305 (same).

142. Note, however, that since nonviolent crimes between 2006 and 2013 contribute both to the postvoluntary period and the pre-presumptive period, we ensure that the post-2013 drop in nonviolent crimes was not used in calculating the postvoluntary period by disallowing the interaction terms to be simultaneously equal to one for the same sentence for nonviolent crimes post-2013. Visually, the differences we estimate can be seen in Figure 2. In this figure, each group and trend are labeled by a number (1)–(9). Our estimate of  $\alpha$  calculates the following for each trend segment:

$$\alpha = E[(3), (2), (8)] - E[(1), (7)] - (E[(6), (5)] - E[(4)])$$

and

$$\beta = E[(9)] - E[(8), (7)] - (E[(6), (3)] - E[(5), (4), (2), (1)])$$

case-specific characteristics, a myriad of indictment, judge, county, circuit, and conviction dummy variables—is also included in some specifications.

We model our outcomes of interest—the total sentence and the imposed sentence, as measured as  $O$  in equation (1)—in a few ways. For one thing, we include as an outcome the raw variable as it occurs in the data. But given the skewed nature of both variables, we also estimate equation (1) with logged outcome variables.<sup>143</sup>

While we are able to observe and control for most of the stated factors that may influence sentence length (for example, severity of crime, prior convictions, and other demographic factors that have been shown to influence sentence lengths like gender<sup>144</sup> and race<sup>145</sup>), there are a host of potentially unobservable variables that may also influence sentence length. To capture these unobserved effects, our model and results include a number of fixed effects including indictment, conviction, judge, circuit, and county fixed effects. Additionally, to capture any sort of statewide changes in trends or sentiment towards sentencing, we also include year or quarter-by-year fixed effects.

Even though these factors are important to fully specifying the model, it is not clear a priori that any of these unobserved factors, if omitted, would necessarily introduce endogeneity or bias to the results. Recall that Alabama made significant legislative changes in introducing voluntary and presumptive sentencing guidelines for some offenses in 2006 and 2013,

---

143. It is also worth noting that, while the vast majority of sentences in our database carry at least some sort of total sentence, around 2% of sentence decisions ended with a zero-sentence due to noncompliant behavior on the part of the court. These occurrences make up about 2% of the data set. An outcome of zero could influence our results more significantly, however, when we consider the months of confinement imposed by the courts. There are a host of reasons why the state's measure of the confinement amounts to zero. In many cases, felons receive prison credit for time spent in jail while awaiting trial, conviction, and sentencing. Additionally, depending on the crime, judges have discretion to substitute confinement with other options that include probation or community corrections. Our data set includes a number of situations where confinement, as opposed to the sentence, comes out to zero. Around 38% of the data set overall falls into this category. In both cases, when logging the data, we employ a similar strategy to that of Professors Rehavi and Starr. See Rehavi & Starr, *supra* note 112, at 1336–37. That is, we add one day to the zeros. Considering the nature of the data set and the circumstances that would lead to zeros in either total sentence length or confinement length, we believe this is a fair assumption. That is, virtually any defendant who received a sentence of any kind almost certainly served at least one day in jail while awaiting bond or trial. To correct for the bias that likely occurs with the OLS estimates of the variance matrix, we cluster the standard errors across two dimensions. First, we cluster at the judge level to account for the idiosyncratic correlation in the errors that undoubtedly occur within each judge's sentencing. Second, since judges are elected by partisan elections, and there is some evidence that judges strategically change behavior relative to election years, we additionally cluster the standard errors at the year level. Carlos Berdejó & Noam Yuchtman, *Crime, Punishment, and Politics: An Analysis of Political Cycles in Criminal Sentencing*, 95 REV. ECON. & STAT. 741, 754 (2013).

144. Sonja B. Starr, *Estimating Gender Disparities in Federal Criminal Cases*, 17 AM. L. & ECON. REV. 127, 154 (2015).

145. Rehavi & Starr, *supra* note 112, at 1349.

respectively.<sup>146</sup> Those dates were decided exogenously by legislators that were relatively removed from the usual daily criminal-sentencing process. In addition to including judge-specific fixed effects, we also conducted a number of tests that leave us reasonably confident that the introduction of sentencing guidelines constituted an exogenous event, meaning that we can reasonably attribute subsequent changes in judicial sentencing behavior to this event.<sup>147</sup>

2. *Measuring the Effect on Racial Sentence Disparities.*—We further this analysis by extending equation (1) to allow for a differential effect of sentencing guidelines by race. We estimate a difference-in-difference-in-

---

146. See *supra* subpart II(A) (describing the changes to Alabama sentencing policy during this time period).

147. For example, even if judges were not expressly involved in the decision-making process of these policy changes, it is still possible that they responded to the changes endogenously for unobserved reasons. We believe that by including judge-specific fixed effects, we mitigate this concern in our model. But to provide an additional layer of confidence, we employ another methodological tool. We examined whether judges attempted to hurry or rush sentencing decisions immediately before the introduction of these policy changes, or conversely, whether judges held out sentencing decisions until after these policy changes. If judges were acting in such a way, we would expect to see a disproportionate amount of sentencing cases decided either the last week of September or the first week of October in 2006 and 2013 relative to the distribution of cases decided all other weeks. To test this, we plotted the distribution of sentencing cases decided by judges each week in our data set relative to the overall distribution. We found that the amount of cases decided during those four critical weeks falls well within the normal flow of weekly case clearances. This bolsters our belief that the introduction of sentencing guidelines was, for all practical purposes, an exogenous event. Another way by which sentence length might be endogenously determined is through manipulation of the charges pursued by prosecutors. For example, Professors Sonja B. Starr and M. Marit Rehavi argued in a 2013 study that much of the literature on the effect of sentencing guidelines on judicial behavior, including seminal reports by the U.S. Sentencing Commission, mistakenly consider a “judge’s final sentencing decision in isolation, ignoring crucial earlier stages of the justice process.” Starr & Rehavi, *supra* note 112, at 5. Unlike many prior studies, Professors Starr and Rehavi found that much of the black–white gap in sentencing actually stemmed from prosecutors’ charging decisions, particularly when prosecutors chose to prosecute individuals with offenses that carried mandatory minimums. *Id.* at 5–6. By factoring in the effect of *Booker* on not just sentencing, but also charging, plea bargaining, and sentencing fact-finding, they failed to find evidence that sentencing discretion actually increased disparity. *Id.* at 5. Thus, critics of our study may argue that any results we identify are not the result of changes in judicial behavior but rather changes in the charging decisions by prosecutors in Alabama operating in the shadow of these new guidelines. This is a reasonable objection. We are able to alleviate this, at least in part but perhaps not completely, by controlling for indictment offense. We are not able, with our data set, to control for the arresting offense, as Professors Starr and Rehavi did in their 2013 study. See *id.* at 24 (describing the usefulness of arrest information in understanding the aggregate sentencing disparities introduced by decisions that may happen from arrest all the way through sentencing). It also bears noting that our identification comes off of the timing of the changes to the law, which is most likely exogenous to the pattern of charges selected by prosecutors. To further this point, we average the severity of the indictment per week (as measured by the indictment score) and find that indictment patterns right before and after each policy change do not stray outside of what would be a weekly average. This was done in a similar fashion to the frequency of sentencing decisions, and corresponding graphs are available upon request. Thus, we feel reasonably confident that our results reflect some genuine changes in judicial behavior, rather than just mere changes in the behavior of prosecutors.

differences, or triple-difference framework, to measure the change in disparity. Essentially, we re-estimate equation (1) by race group and compare the results. Formally, this is estimated by interacting each time and group variable in equation (1) with the race variable. The result is the following model, which we describe as equation (2):

$$\begin{aligned}
 O_{it} = & a_1 \text{Group}V_{it} + a_2 \text{Post}V_{it} + \alpha \text{Group}V_{it} * \text{Post}V_{it} + b_1 \text{Group}P_{it} \\
 & + b_2 \text{Post}P_{it} + \beta \text{Group}P_{it} * \text{Post}P_{it} + c_1 \text{black}_{it} \\
 & + c_2 \text{Group}V_{it} * \text{black}_{it} + c_3 \text{Post}V_{it} * \text{black}_{it} + c_4 \text{Group}V_{it} \\
 & * \text{Post}V_{it} * \text{black}_{it} + c_5 \text{Group}P_{it} * \text{black}_{it} + c_6 \text{Post}P_{it} \\
 & * \text{black}_{it} + c_7 \text{Group}P_{it} * \text{Post}P_{it} * \text{black}_{it} + X_{it} + \varepsilon_{it}
 \end{aligned} \tag{2}$$

Much of equation (2) mirrors that of equation (1), except the coefficients of interest are the triple interactions of coefficients  $c_4$  for the effect of the voluntary worksheet and  $c_7$  for the presumptive worksheet. For instance, the interpretation of  $c_7$  is an estimate of the change in relative disparity in race where a negative coefficient suggests a decrease in the racial gap and a positive coefficient suggests an increase in the racial gap.

3. *Measuring the Effect on Interjudge Sentencing Disparities.*—To further address the effect these sentencing changes had on sentence lengths, we extend the difference-in-differences methodology explained previously to measure more directly the effect on interjudge disparities after the implementation of voluntary and presumptive sentencing guidelines. To do this, we develop a triple-differences methodology. This allows us to directly measure the degree to which the gap in sentence lengths closed between the most-extreme sentencing behaviors on both ends of the spectrum. We essentially calculate the difference-in-differences estimator separately for both the most lenient and the toughest judges and then difference those differences. Recall that our main result from equation (1) is calculated as:

$$\beta = [O_{t \geq 2013, P} - O_{t < 2013, P}] - [O_{t \geq 2013, C} - O_{t < 2013, C}] \tag{3}$$

where  $O_{t \geq 2013, P}$  is the outcome for the treated group,  $P$ , after the presumptive guidelines came into effect ( $t \geq 2013$ );  $O_{t < 2013, P}$  is that same group before the guidelines became presumptive; and  $[O_{t \geq 2013, C} - O_{t < 2013, C}]$  is the same difference across time for the control group,  $C$ . The triple difference calculates  $\beta$  separately for the most lenient and strictest judges, and then differences those two effects, or:

$$\beta^{\text{toughest}} - \beta^{\text{most lenient}}$$

which expands to:

$$\left\{ \left[ O_{t \geq 2013, P}^{toughest} - O_{t < 2013, P}^{toughest} \right] - \left[ O_{t \geq 2013, C}^{toughest} - O_{t < 2013, C}^{toughest} \right] \right\} - \left\{ \left[ O_{t \geq 2013, P}^{most\ lenient} - O_{t < 2013, P}^{most\ lenient} \right] - \left[ O_{t \geq 2013, C}^{most\ lenient} - O_{t < 2013, C}^{most\ lenient} \right] \right\}^{148} \quad (4)$$

What results is an estimate that speaks to the degree to which the gap between the two polar-opposite ends of the sentencing spectrum converge. As we explain in more detail in the next Part, we find the gap in total sentence length between the two extreme types of judges closed by about ten months—which is statistically significant—when the worksheet became presumptive. But we find no evidence of such an effect when it became voluntary.<sup>149</sup>

### III. Findings

Overall, we find compelling evidence that the introduction of sentencing guidelines in Alabama contributed to reductions in sentence length, reductions in racial disparities in sentences for similar offenses, and reductions in interjudge disparities in sentence lengths. While voluntary guidelines may have had modest effects, we find stronger evidence that presumptive sentencing guidelines contributed to these outcomes. These results are highly significant, even when controlling for alternative possible explanations.

Further, we find evidence that the introduction of the sentencing guidelines did not affect all judges equally. Voluntary guidelines appear reasonably effective at altering the behavior of judges who fall within the middle quartile in previous sentence lengths. But voluntary guidelines appear to do little to rein in the behavior of judges with a history of particularly punitive sentencing histories relative to their peers in the judiciary. By contrast, presumptive sentencing guidelines appear more effective at altering the behavior of these historically tough judges.

---

148. Econometrically, this is executed much in the same way as equation (1) except that every key difference-in-differences variable is fully interacted with the set of “toughness” dummy variables. Given the nature of the triple differences technique, we are unable to include the entire set of different combinations of fixed effects in our result tables. It is worth noting, though, that our results are robust to the inclusion or exclusion of various variables.

149. It is important not to rely entirely on the raw data in making this assessment. The toughest judges sentenced applicable drug and nonviolent property crimes prior to the 2013 change at an average of 101 months, whereas the most lenient judges only sixty-four months. After the worksheet became presumptive, the average sentence length for the toughest judges dropped by nearly forty months to sixty-four, while the most lenient judges dropped only eight months to an average sentence length of fifty-six months. Thus, prior to the change they were about forty months apart, and after the change they were only eight months apart. This sort of raw analysis does not, however, account for general trends in sentencing, which we capture in the second set of differences that we calculate from the nonaffected crimes—which for the toughest judges also dropped by about fourteen months.



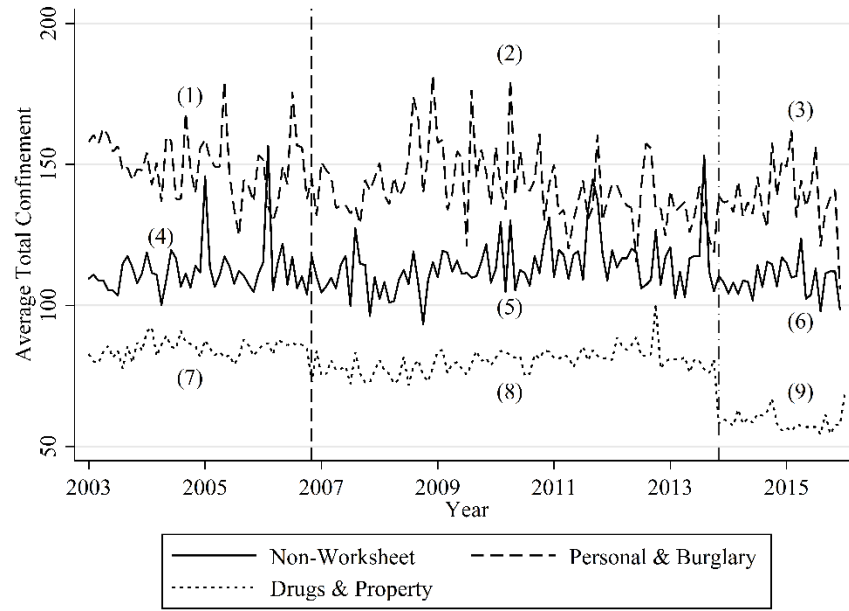
Overall, these differing effects on judicial behavior have real impacts on the average length of sentences handed down by trial judges. We find that the introduction of voluntary sentencing guidelines in Alabama reduced the average sentence length by around seven months. But when the guidelines became presumptive, the average sentence length dropped by almost two years.

#### A. *Trends in Raw Data*

As a preliminary matter, it may be helpful to start our discussion by looking at trends in the raw data. While our formal difference-in-differences and triple-differences regressions can provide more nuanced explanations, much of the story can be told by looking at this sort of raw data.

First, the raw data show that presumptive sentencing guidelines appear to have exerted a sudden and strong downward influence on the length of confinement for drug and personal-property offenses. There was no similar reduction in overall sentence length when Alabama initially made these sentences merely advisory. Figure 2 visually illustrates the length of confinement over time for the three different categories of offenses described above in subpart II(A): (1) *nonworksheet offenses* that were never subject to any sort of voluntary or presumptive sentencing guidelines over this time period; (2) *personal and burglary offenses* that were subject to voluntary sentencing guidelines starting in October 2006 through the end of our timeline; and (3) *drug and property offenses* that were subject to voluntary sentencing guidelines in October 2006 and presumptive sentencing guidelines from October 2013 through the end of our timeline. The vertical lines in the figure signal when Alabama introduced voluntary and presumptive sentencing guidelines for certain offenses in late 2006 and 2013.

Figure 2: Trends in Confinement for Each Offense Category Over Time

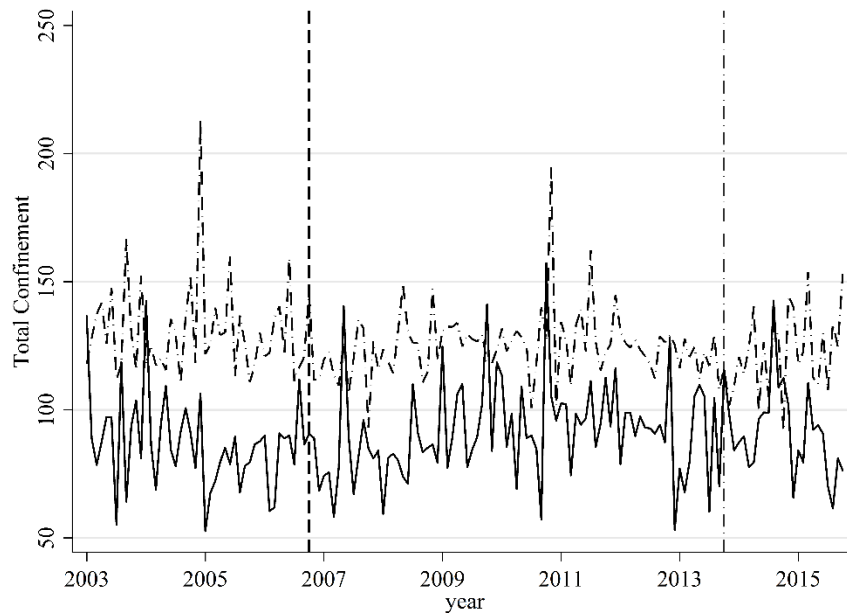


As expected, there was no obvious change in the overall period of confinement for these nonworksheet offenses over time. This finding makes intuitive sense. Since these offenses were not subject to any of the sentencing guidelines discussed in this study, we would expect total confinement for these offenses to stay relatively stable over time. Similarly, it appears that the introduction of voluntary sentencing guidelines in 2006 for personal and burglary offenses did not result in any immediately noticeable effects on overall confinement. By contrast, we see a significant shift in the total confinement for drug and property offenses starting in October 2013. Again, like personal and burglary offenses, these crimes were subject to voluntary sentencing guidelines in October 2006. And like personal and burglary offenses, the introduction of voluntary sentencing guidelines appeared to have had a modest effect on overall confinement periods. But almost immediately after the introduction of presumptive sentencing guidelines in October 2013, the average period of confinement handed down by Alabama judges dropped visibly. Thus, the raw data suggest that presumptive sentencing guidelines likely exerted a more significant downward effect on overall sentence length than voluntary sentencing guidelines, which appear to have had a smaller overall effect on sentence lengths.

Second, the raw data suggest that the introduction of presumptive sentencing guidelines contributed to a reduction in disparities between the historically harshest and most lenient judges in Alabama. We find minimal evidence in the raw data to suggest that the voluntary guidelines correlated

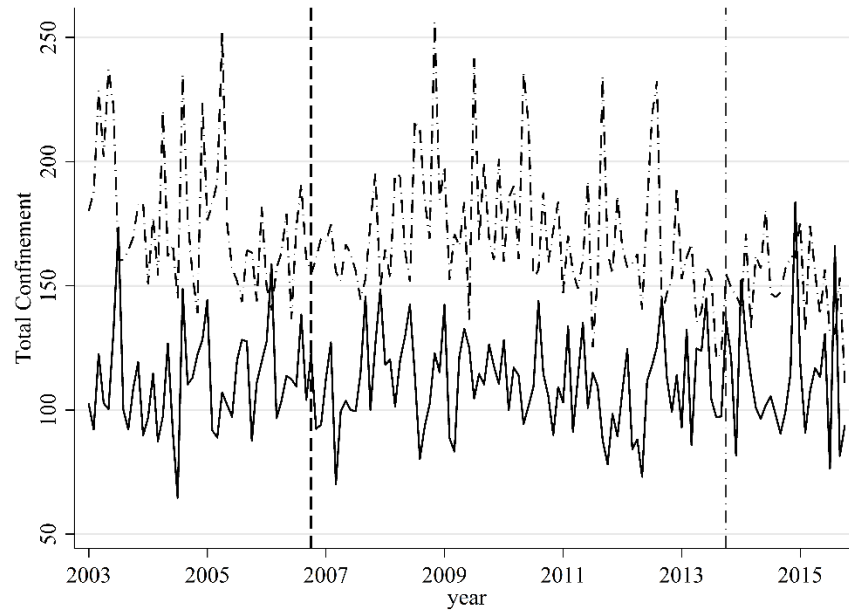
with a similar reduction in interjudge sentence disparities. Figures 3 through 5 present data similar to Figure 2. That is, they show trends in the average total confinement for defendants over time. But rather than lumping together all defendants into one category, Figures 3 through 5 divide these cases into two categories: sentences handed down by the harshest quartile of Alabama judges (represented by the dashed lines) and those handed down by the most lenient quartile of judges (represented by the solid line). Thus, the space between these two trend lines represents the level of disparities between the harshest and most lenient judges. A large gap between these two trend lines represents a large interjudge disparity. A small gap between these two trend lines represents a small interjudge disparity. Again, each figure has vertical lines signifying the dates that the state began introducing some voluntary (2006) and presumptive (2013) guidelines.

*Figure 3: Inter-Judge Disparities in Sentences for Non-Worksheet Offenses*



Remember, nonworksheet offenses have never been subject to voluntary or presumptive sentencing guidelines in Alabama over this time period. Thus, they serve as a sort of control for our study. And as expected, we see little evidence of a change in sentencing disparities for these nonworksheet offenses over this time period.

*Figure 4: Inter-Judge Disparities in Sentences for Personal and Burglary Offenses*



By contrast, personal and burglary offenses were subject to the voluntary sentencing worksheet starting in 2006 through the end of our timeline. Based on the trends in raw data, it does not appear that the introduction of voluntary sentences exerted a dramatic effect on the disparities between the harshest and most lenient judges in Alabama. The space between the two trend lines may have narrowed slightly, but it remains relatively stable throughout this time period.

*Figure 5: Inter-Judge Disparities in Sentences for Drugs and Property Offenses*

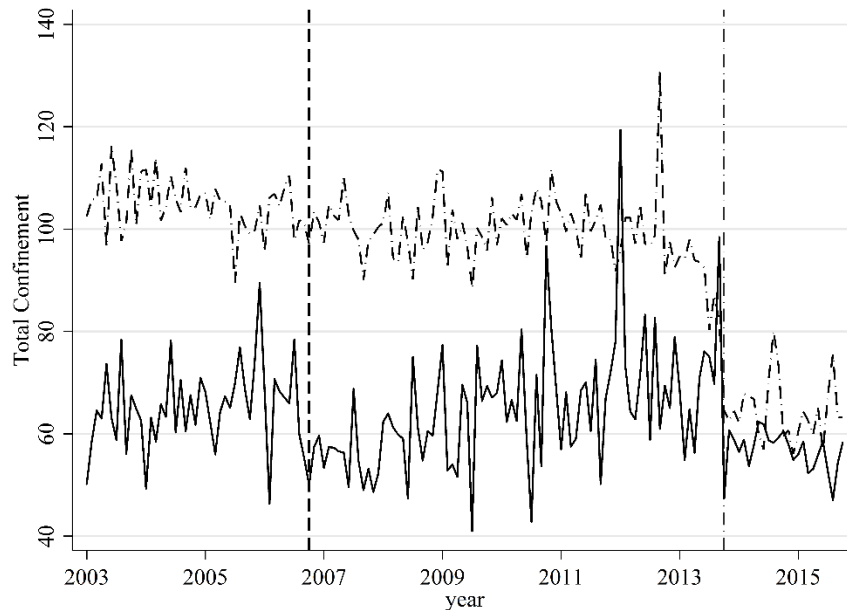
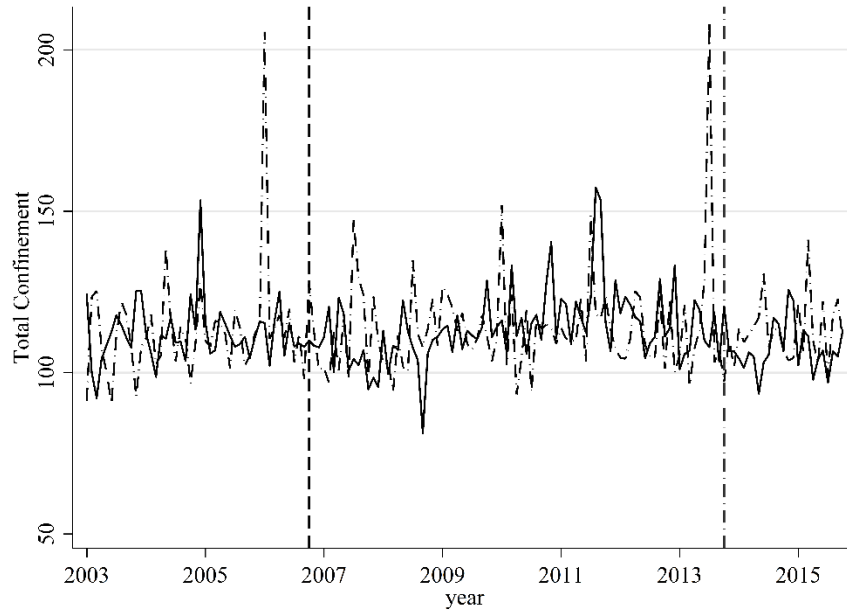


Figure 5, though, shows a remarkable drop in the apparent interjudge sentencing disparities when Alabama introduced presumptive sentencing guidelines for drug and personal-property offenses. This appears to be the result of the harshest judges in Alabama substantially reducing penalties to mirror more closely those given by the more lenient judges in the state. All of this suggests that, to the extent sentencing guidelines affected interjudge sentencing disparities, presumptive sentencing guidelines may be doing most of the work.

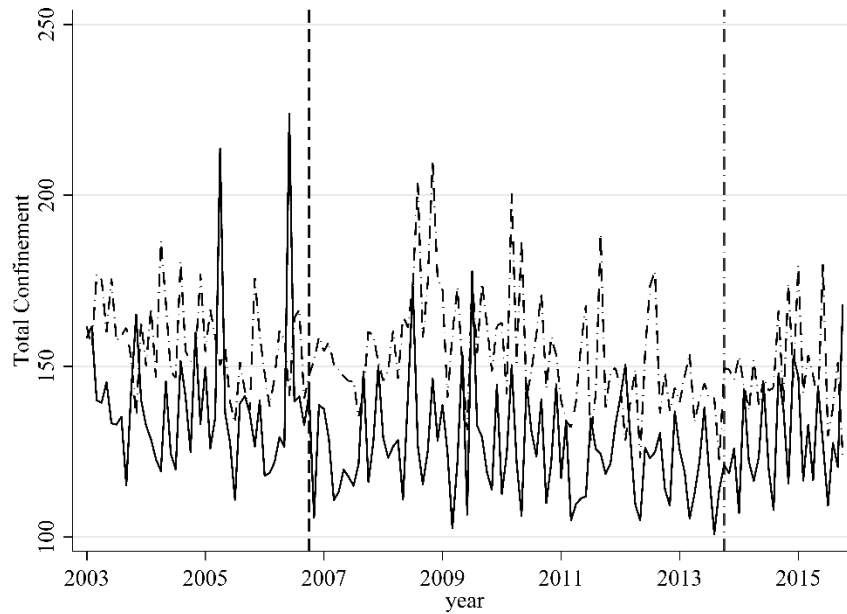
Third, the raw data suggest that the introduction of sentencing guidelines may have contributed to reductions in racial disparities in sentences. Figures 6 through 8 present the same data as Figure 2, only broken down by race. Thus, when these figures show one race receiving longer terms of confinement than another race, this is suggestive of possible racial disparities. In these figures, the solid lines represent sentences given to white defendants, while the dashed lines represent sentences given to black defendants.

*Figure 6: Racial Disparities in Sentences for Non-Worksheet Offenses*



There appears to be minimal evidence in the raw data to suggest racial bias in nonworksheet offenses, which are not subject to any of the sentencing guidelines described in this study over this time period. And as we would expect, there is no relationship between the introduction of sentencing guidelines for other offenses and racial disparities in total confinement for these nonworksheet offenses. Sentences remain relatively stable over time, regardless of race.

*Figure 7: Racial Disparities in Sentences for Personal and Burglary Offenses*



Similarly, there does not appear to be strong evidence in the raw data to suggest that the introduction of voluntary sentencing guidelines for personal and burglary offenses contributed to any significant changes in racial disparities. While there is some evidence that black defendants may have received somewhat higher sentences on average than white defendants over our time period, the introduction of voluntary sentencing guidelines does not appear, at least from the raw data, to exert an effect on racial disparities for this subset of offenses. And overall, the raw data do not paint a picture of significant racial disparities in sentences within this category of offenses.

*Figure 8: Racial Disparities in Sentences for Drug and Property Offenses*

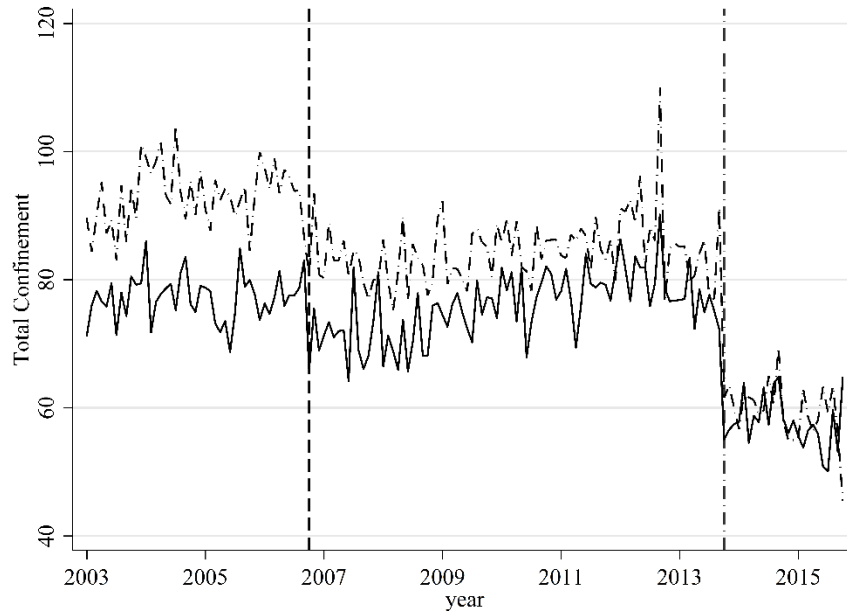


Figure 8 tells a very different story. Before the introduction of voluntary sentencing guidelines in October 2006, there is stronger evidence in the raw data of possible racial disparities in sentence lengths. Black defendants, represented by the dashed line in Figure 8, frequently seem to receive somewhat harsher sentences than their white counterparts. It appears that the introduction of voluntary sentencing guidelines in 2006 may have reduced this racial disparity somewhat. But the raw data suggest that the introduction of presumptive sentencing guidelines virtually eliminated any apparent evidence of racial disparities.

Combined, these preliminary raw data suggest that sentencing guidelines—particularly presumptive sentencing guidelines—may have a significant effect on judicial behavior. The introduction of presumptive sentencing guidelines correlates with a noticeable reduction in overall sentence length, interjudge disparities, and racial disparities. Nevertheless, these raw data leave many questions unanswered. These raw data do not control for other variables that may be influencing sentence lengths and disparities. The regressions below introduce additional controls to address this problem. Additionally, these raw data provide only a small amount of information on the behavior of individual judges. Are all judges altering their behavior equally in response to voluntary and presumptive sentencing guidelines? Or do these guidelines influence judges differently depending on the judge's underlying characteristics? Answering these questions may provide important insight into the causal mechanism behind the apparent



effectiveness of these sentencing guidelines in influencing judicial behavior. Answering these questions requires a more sophisticated model, as previously described above in subpart II(C). The next two subparts describe the results of this modeling.

### *B. Reduction in Sentence Lengths*

Our first model uses a more sophisticated difference-in-differences approach to estimate the causal effect of voluntary and presumptive sentencing guidelines on total sentence lengths given by trial judges in Alabama. The resulting estimates of  $\alpha$  and  $\beta$  for our four outcomes,  $O$ , can be found in Figures 9 through 12. Each table reports ten different specifications of the same outcome. Column (1) of each table reports the straight difference-in-differences estimate without any other controls or fixed effects. Column (2) in each table introduces case-specific controls mentioned previously in subpart II(C) and year-fixed effects. Columns (3) through (10) introduce various arrays of judge, circuit, county, offense, and year effects, and again, we report two-way clustered standard errors clustered across judge and year.

We see in Figures 9 and 10 a persistent story regardless of whether we measure total sentence length in levels or logs. That is, the imposition of presumptive sentencing guidelines is associated with a statistically significant reduction in total sentence lengths. Across all specifications in Figure 9, we estimated that Alabama's move to the presumptive worksheet contributed to a twenty-four-month reduction in the average total sentence for nonviolent drug and property crimes, or in Figure 10, between a 31% and 44% reduction in total sentence. We see a similar, albeit smaller, effect for the imposition of the voluntary sentencing guidelines. After Alabama implemented voluntary worksheets, judges issued sentences that were an average of between six and seven months, or between 8% and 18%, shorter.

Figure 9: Effect of Voluntary and Presumptive Sentence Guidelines on Total Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Presumptive Guidelines	-24.30 <sup>‡</sup> (2.03)	-24.78 <sup>‡</sup> (1.97)	-25.48 <sup>‡</sup> (2.81)	-25.54 <sup>‡</sup> (2.82)	-25.35 <sup>‡</sup> (2.82)	-25.41 <sup>‡</sup> (2.83)	-25.10 <sup>‡</sup> (3.00)	-25.17 <sup>‡</sup> (3.01)	-24.98 <sup>‡</sup> (3.00)	-25.06 <sup>‡</sup> (3.01)
Voluntary Guidelines	-7.24 <sup>‡</sup> (2.06)	-8.13 <sup>‡</sup> (2.16)	-7.65 <sup>‡</sup> (2.65)	-7.64 <sup>‡</sup> (2.64)	-7.68 <sup>‡</sup> (2.64)	-7.67 <sup>‡</sup> (2.64)	-6.71 <sup>‡</sup> (2.50)	-6.69 <sup>‡</sup> (2.53)	-6.74 <sup>‡</sup> (2.50)	-6.73 <sup>‡</sup> (2.53)
Sample Size	221,894	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348
R-squared	0.05	0.15	0.21	0.21	0.21	0.21	0.22	0.22	0.22	0.22
Base Controls		X	X	X	X	X	X	X	X	X
<i>Fixed Effects</i>										
Circuit			X	X			X	X		X
County					X	X			X	X
Judge			X	X	X	X	X	X	X	X
Most Serious Indictment			X	X	X	X	X	X	X	X
Most Serious Conviction							X	X	X	X
Year Effects		X	X		X		X		X	
Quarter by Year Effects				X		X		X		X

Notes: Each observation is measured at the individual case level. Each column represents a unique regression, and standard errors are clustered at the judge and year level. The outcome of interest is the total sentence length in months and the base controls consist of the most serious indictment/conviction score, plea deal dummy, indictment counts, seriousness of priors, penitentiary ruling, split sentence dummy, male, black, alcohol anonymous ruling, drug court or mental health court, habitual offender flag, and drugs near schools or projects flag.

<sup>^</sup>  $p < 0.10$ , <sup>†</sup>  $p < 0.05$ , <sup>‡</sup>  $p < 0.01$

Figure 10: Effect of Voluntary and Presumptive Sentence Guidelines on the Log of Total Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Presumptive Guidelines	-0.44 <sup>‡</sup> (0.03)	-0.39 <sup>‡</sup> (0.03)	-0.32 <sup>‡</sup> (0.04)	-0.32 <sup>‡</sup> (0.04)	-0.32 <sup>‡</sup> (0.04)	-0.32 <sup>‡</sup> (0.04)	-0.31 <sup>‡</sup> (0.04)	-0.31 <sup>‡</sup> (0.04)	-0.31 <sup>‡</sup> (0.04)	-0.31 <sup>‡</sup> (0.04)
Voluntary Guidelines	-0.18 <sup>‡</sup> (0.03)	-0.16 <sup>‡</sup> (0.04)	-0.09 <sup>‡</sup> (0.03)	-0.09 <sup>‡</sup> (0.03)	-0.09 <sup>‡</sup> (0.03)	-0.09 <sup>‡</sup> (0.03)	-0.08 <sup>‡</sup> (0.03)	-0.08 <sup>‡</sup> (0.03)	-0.08 <sup>‡</sup> (0.03)	-0.08 <sup>‡</sup> (0.03)
Sample Size	221,894	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348
R-squared	0.02	0.15	0.26	0.26	0.27	0.27	0.28	0.28	0.29	0.29
Base Controls		X	X	X	X	X	X	X	X	X
<i>Fixed Effects</i>										
Circuit			X	X			X	X		
County					X	X			X	X
Judge			X	X	X	X	X	X	X	X
Most Serious Indictment			X	X	X	X	X	X	X	X
Most Serious Conviction							X	X	X	X
Year Effects		X	X		X		X		X	
Quarter by Year Effects				X		X		X		X

Notes: Each observation is measured at the individual case level. Each column represents a unique regression, and standard errors are two way clustered at the judge and year level. The outcome of interest is the total sentence length in months and the base controls consist of the most serious indictment/conviction score, plea deal dummy, indictment counts, seriousness of priors, penitentiary ruling, split sentence dummy, male, black, alcohol anonymous ruling, drug court or mental health court, habitual offender flag, and drugs near schools or projects flag.

<sup>^</sup>  $p < 0.10$ , <sup>†</sup>  $p < 0.05$ , <sup>‡</sup>  $p < 0.01$

Given the sentencing dynamics specific to Alabama, we predictably see similar but smaller-in-magnitude reductions in the sentences imposed to offenders.<sup>150</sup> Figure 11 displays the results on the sentence imposed in raw terms, and Figure 12 displays the results as a logged outcome with the results presented as percentage changes. Recalling the average differences in base sentencing between total and imposed sentences, we observe a smaller drop in the imposed sentence compared to the total sentence, but the drops are proportionally similar in magnitude—around 30%. It is important to note, however, that the evidence is much less clear that the voluntary worksheet had any impact on sentences imposed, as many of the results in Figures 11 and 12 are statistically insignificant.

---

150. While the total sentence represents the threat of incarceration if the sentencing terms are not met by the offender, recall that the sentence imposed represents an approximation of the time actually served.

Figure 11: Effect of Voluntary and Presumptive Sentence Guidelines on the Imposed Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Presumptive Guidelines	-2.15 (1.80)	-6.75‡ (1.60)	-7.97‡ (2.42)	-7.98‡ (2.42)	-7.95‡ (2.46)	-7.97‡ (2.45)	-7.20‡ (2.53)	-7.23‡ (2.52)	-7.19‡ (2.55)	-7.22‡ (2.55)
Voluntary Guidelines	-0.04 (1.66)	-2.92^ (1.54)	-3.45^ (1.98)	-3.45^ (1.97)	-3.54^ (1.95)	-3.54^ (1.94)	-2.22 (1.85)	-2.21 (1.84)	-2.31 (1.83)	-2.30 (1.82)
Sample Size	221,894	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348
R-squared	0.03	0.13	0.19	0.19	0.19	0.20	0.21	0.21	0.21	0.21
Base Controls		X	X	X	X	X	X	X	X	X
<i>Fixed Effects</i>										
Circuit			X	X			X	X		
County					X	X			X	X
Judge			X	X	X	X	X	X	X	X
Most Serious Indictment			X	X	X	X	X	X	X	X
Most Serious Conviction							X	X	X	X
Year Effects		X	X		X		X		X	
Quarter by Year Effects				X		X		X		X

Notes: Each observation is measured at the individual case level. Each column represents a unique regression, and standard errors are clustered at the judge and year level. The outcome of interest is the total sentence length in months and the base controls consist of the most serious indictment/conviction score, plea deal dummy, indictment counts, seriousness of priors, penitentiary ruling, split sentence dummy, male, black, alcohol anonymous ruling, drug court or mental health court, habitual offender flag, and drugs near schools or projects flag.

^  $p < 0.10$ , †  $p < 0.05$ , ‡  $p < 0.01$

Figure 12: Effect of Voluntary and Presumptive Sentence Guidelines on the Log of the Imposed Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Presumptive Guidelines	-0.32‡ (0.11)	-0.27‡ (0.06)	-0.26‡ (0.09)	-0.26‡ (0.09)	-0.27‡ (0.09)	-0.27‡ (0.09)	-0.25‡ (0.09)	-0.25‡ (0.09)	-0.25‡ (0.09)	-0.26‡ (0.09)
Voluntary Guidelines	-0.21‡ (0.09)	-0.15‡ (0.06)	-0.14‡ (0.07)	-0.14‡ (0.07)	-0.14‡ (0.07)	-0.14‡ (0.07)	-0.12 <sup>^</sup> (0.07)	-0.12 <sup>^</sup> (0.07)	-0.13‡ (0.07)	-0.13‡ (0.07)
Sample Size	221,894	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348
R-squared	0.04	0.36	0.44	0.45	0.45	0.45	0.45	0.45	0.45	0.45
Base Controls		X	X	X	X	X	X	X	X	X
<i>Fixed Effects</i>										
Circuit			X	X			X	X		X
County					X	X			X	X
Judge			X	X	X	X	X	X	X	X
Most Serious Indictment			X	X	X	X	X	X	X	X
Most Serious Conviction							X	X	X	X
Year Effects		X	X		X		X		X	
Quarter by Year Effects				X		X		X		X

Notes: Each observation is measured at the individual case level. Each column represents a unique regression, and standard errors are clustered at the judge and year level. The outcome of interest is the total sentence length in months and the base controls consist of the most serious indictment/conviction score, plea deal dummy, indictment counts, seriousness of priors, penitentiary ruling, split sentence dummy, male, black, alcohol anonymous ruling, drug court or mental health court, habitual offender flag, and drugs near schools or projects flag.

<sup>^</sup>  $p < 0.10$ , <sup>†</sup>  $p < 0.05$ , <sup>‡</sup>  $p < 0.01$

In total, there is strong evidence to suggest that the decision by Alabama legislators to make the sentencing worksheet presumptive had a statistically significant effect on the total sentence and sentence imposed. There is also some evidence to suggest the introduction of voluntary guidelines may have had a smaller but still significant effect on total sentences. Nevertheless, there is less evidence to suggest that the voluntary worksheet has any effect on the sentence imposed.

Perhaps unsurprisingly, our results lead us to conclude that, while some judges respond to the imposition of voluntary sentencing guidelines, presumptive guidelines are more effective at altering judicial behavior and reducing sentence length. But this leaves open important causal questions: Why are presumptive sentencing guidelines more effective, in the aggregate, in altering sentencing behavior? Do presumptive sentencing guidelines force changes in behavior among a different class of judges than voluntary guidelines? The next subparts help answer these questions.

### *C. Reductions in Racial Disparities*

Recall that Figures 6 through 8 showed that the limited evidence for possible racial disparities in sentencing in Alabama during this time frame appeared in the case of drug and property offenses. For these offenses, it appeared that black defendants received somewhat longer sentences than their white counterparts. But after the implementation of presumptive sentencing guidelines, it appeared that this gap between sentences for white and black defendants mostly dissipated.

Formally as seen in Figure 13, our triple-difference estimation strategy confirms these observations in the raw data. Presumptive sentencing guidelines appeared to close the racial gap in disparities by about eight months from the total sentence. This result holds irrespective of whether total sentence is measured at the level or log, as seen in Figure 14. There is, however, no evidence that the voluntary sentencing guidelines had any statistically significant effect on closing racial disparities. Additionally, as seen in Figures 15 and 16, there is less compelling evidence of narrowing of the racial disparity gap in imposed sentences.

Figure 13: Effect of Voluntary and Presumptive Sentence Guidelines on Racial Disparities in Total Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Presumptive Guidelines	-10.09 <sup>‡</sup> (4.13)	-12.54 <sup>‡</sup> (4.32)	-7.92 <sup>^</sup> (4.72)	-7.98 <sup>^</sup> (4.71)	-7.86 <sup>^</sup> (4.74)	-7.91 <sup>^</sup> (4.74)	-8.81 <sup>^</sup> (4.86)	-8.85 <sup>^</sup> (4.88)	-8.75 <sup>^</sup> (4.86)	-8.79 <sup>^</sup> (4.89)
Voluntary Guidelines	-0.03 (2.99)	-3.40 (3.39)	1.01 (4.15)	1.03 (4.11)	0.93 (4.16)	0.95 (4.12)	-0.32 (4.21)	-0.28 (4.18)	-0.42 (4.22)	-0.38 (4.19)
Sample Size	221,894	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348
R-squared	0.05	0.15	0.21	0.21	0.21	0.21	0.22	0.22	0.22	0.22
Base Controls		X	X	X	X	X	X	X	X	X
<i>Fixed Effects</i>										
Circuit			X	X			X	X		
County					X	X			X	X
Judge			X	X	X	X	X	X	X	X
Indictment Offense			X	X	X	X	X	X	X	X
Conviction Offense							X	X	X	X
Year Effects		X	X		X		X		X	
Quarter by Year Effects				X		X		X		X

<sup>^</sup>  $p < 0.10$ , <sup>†</sup>  $p < 0.05$ , <sup>‡</sup>  $p < 0.01$



Figure 14: Effect of Voluntary and Presumptive Sentence Guidelines on Racial Disparities in Log of Total Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Presumptive Guidelines	-0.22‡ (0.04)	-0.26‡ (0.08)	-0.23‡ (0.07)	-0.23‡ (0.07)	-0.22‡ (0.07)	-0.22‡ (0.07)	-0.20‡ (0.07)	-0.20‡ (0.07)	-0.19‡ (0.07)	-0.19‡ (0.07)
Voluntary Guidelines	-0.10‡ (0.03)	-0.14^ (0.08)	-0.09^ (0.05)	-0.09^ (0.05)	-0.09^ (0.05)	-0.09^ (0.05)	-0.08 (0.05)	-0.08 (0.05)	-0.08 (0.05)	-0.08 (0.05)
Sample Size	221,894	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348
R-squared	0.02	0.15	0.26	0.26	0.27	0.27	0.28	0.28	0.29	0.29
Base Controls		X	X	X	X	X	X	X	X	X
<i>Fixed Effects</i>										
Circuit			X	X			X	X		
County					X	X			X	X
Judge			X	X	X	X	X	X	X	X
Most Serious Indictment Offense			X	X	X	X	X	X	X	X
Most Serious Conviction Offense							X	X	X	X
Year Effects		X	X	X	X	X	X	X	X	X
Quarter by Year Effects				X	X	X	X	X	X	X

^  $p < 0.10$ , †  $p < 0.05$ , ‡  $p < 0.01$

Figure 15: Effect of Voluntary and Presumptive Sentence Guidelines on Racial Disparities in the Imposed Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Presumptive Guidelines	1.20 (3.47)	-1.55 (4.02)	-0.07 (4.21)	-0.08 (4.20)	-0.05 (4.22)	-0.06 (4.21)	-0.98 (4.39)	-0.98 (4.37)	-0.93 (4.38)	-0.92 (4.36)
Voluntary Guidelines	2.33 (3.45)	-0.60 (3.68)	0.41 (3.97)	0.41 (3.96)	0.44 (3.97)	0.44 (3.96)	-0.66 (4.15)	-0.65 (4.11)	-0.63 (4.13)	-0.61 (4.10)
Sample Size	221,894	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348
R-squared	0.03	0.13	0.19	0.19	0.19	0.20	0.21	0.21	0.21	0.21
Base Controls		X	X	X	X	X	X	X	X	X
<i>Fixed Effects</i>										
Circuit			X	X			X	X		
County					X	X			X	X
Judge			X	X	X	X	X	X	X	X
Most Serious Indictment Offense			X	X	X	X	X	X	X	X
Most Serious Conviction Offense							X	X	X	X
Year Effects		X	X		X		X		X	
Quarter by Year Effects				X		X		X		X

<sup>^</sup>  $p < 0.10$ , <sup>†</sup>  $p < 0.05$ , <sup>‡</sup>  $p < 0.01$

Figure 16: Effect of Voluntary and Presumptive Sentence Guidelines on the Racial Disparity in Log of the Imposed Sentence Length

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Presumptive Guidelines	-0.32‡ (0.11)	-0.27‡ (0.06)	-0.26‡ (0.09)	-0.26‡ (0.09)	-0.27‡ (0.09)	-0.27‡ (0.09)	-0.25‡ (0.09)	-0.25‡ (0.09)	-0.25‡ (0.09)	-0.26‡ (0.09)
Voluntary Guidelines	-0.21‡ (0.09)	-0.15‡ (0.06)	-0.14‡ (0.07)	-0.14‡ (0.07)	-0.14‡ (0.07)	-0.14‡ (0.07)	-0.12‡ (0.07)	-0.12 <sup>^</sup> (0.07)	-0.13‡ (0.07)	-0.13‡ (0.07)
Sample Size	221,894	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348	221,348
R-squared	0.04	0.36	0.44	0.45	0.45	0.45	0.45	0.45	0.45	0.45
Base Controls		X	X	X	X	X	X	X	X	X
<i>Fixed Effects</i>										
Circuit			X	X			X	X		
County					X	X			X	X
Judge			X	X	X	X	X	X	X	X
Most Serious Indictment Offense			X	X	X	X	X	X	X	X
Most Serious Conviction Offense							X	X	X	X
Year Effects		X	X		X		X		X	
Quarter by Year Effects				X		X		X		X

<sup>^</sup>  $p < 0.10$ , <sup>†</sup>  $p < 0.05$ , <sup>‡</sup>  $p < 0.01$

This lack of a consistent finding as to the effect of sentencing guidelines on the length of imposed sentences is a bit puzzling. This may be, though, due to a lack of preexisting disparities. Or it may be due to the fact that the imposed sentence length is more easily amenable to other forms of bias that cannot be fully addressed by the sentencing guidelines. Overall, though, we find fairly compelling evidence that voluntary and presumptive sentences had several significant effects on judicial behavior.

*D. Reductions in Interjudge Disparities*

Perhaps unsurprisingly, our results thus far lead us to conclude that, while some judges respond to the imposition of voluntary sentencing guidelines, presumptive guidelines are more effective at altering judicial behavior and reducing sentence length. But why are presumptive sentencing guidelines more effective, in the aggregate, in altering sentencing behavior? Do presumptive sentencing guidelines force changes in behavior among a different class of judges than voluntary guidelines?

To address these questions, we parse our estimates of interest ( $\alpha$  and  $\beta$  from equation (1)) into four separate variables for each policy change. Using the sentencing data prior to the first policy change, we create a variable that places each judge in quartiles of sentencing “toughness” and interact that with our difference-in-differences variables. Essentially, we are attempting to see if judges who were tougher prior to the policy changes reacted differently than judges with a history of leniency in sentencing. Formally, for instance, our estimate of  $\beta$  from equation (1) becomes:

$$\beta = \beta^{most\ lenient} + \beta^{lighter} + \beta^{tougher} + \beta^{toughest} \quad (5)$$

We report this result in two separate figures: Figure 17 for the voluntary coefficients and Figure 18 for the presumptive coefficients, keeping in mind that both figures report results from the same regression.<sup>151</sup> The judges in the lowest quartile of sentencing are those judges who, prior to the policy changes, issued the most lenient sentences for similarly situated defendants relative to their peers in the state. By contrast, those in the highest quartile were the toughest relative to their peers on similarly situated defendants. And those in the middle quartiles fell somewhere between the toughest and most lenient of judges in the state before these changes.

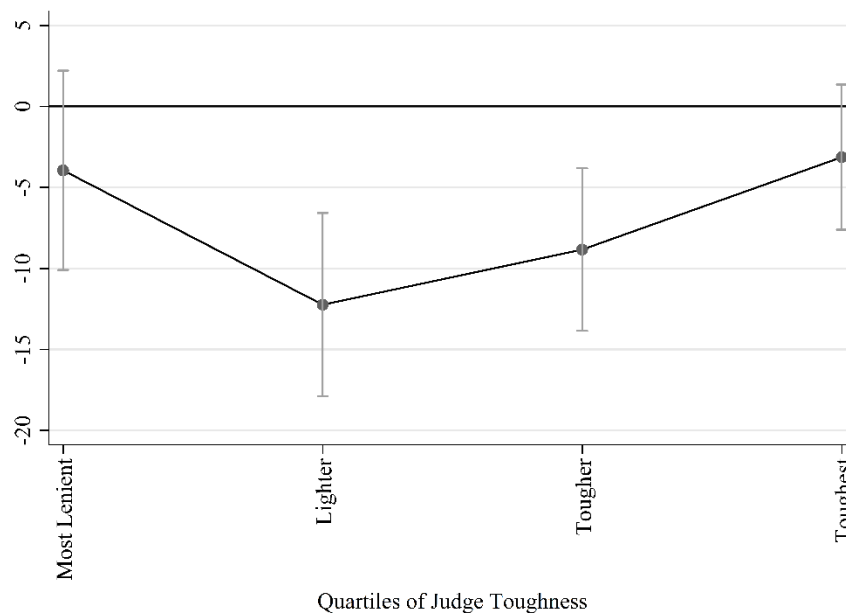
Figure 17 shows the effect of voluntary sentencing guidelines on the length of sentences imposed after the imposition of voluntary guidelines.

---

151. Since the aim of this analysis is to identify judge-specific heterogeneities, we report only the results for total sentencing because, based on the previous results, it appears that judges responded most starkly in total sentencing decisions. For ease of presentation, we report only the results that include fixed effects for circuit, judge, year, and indictment, though the results are robust to other combinations of fixed effects.

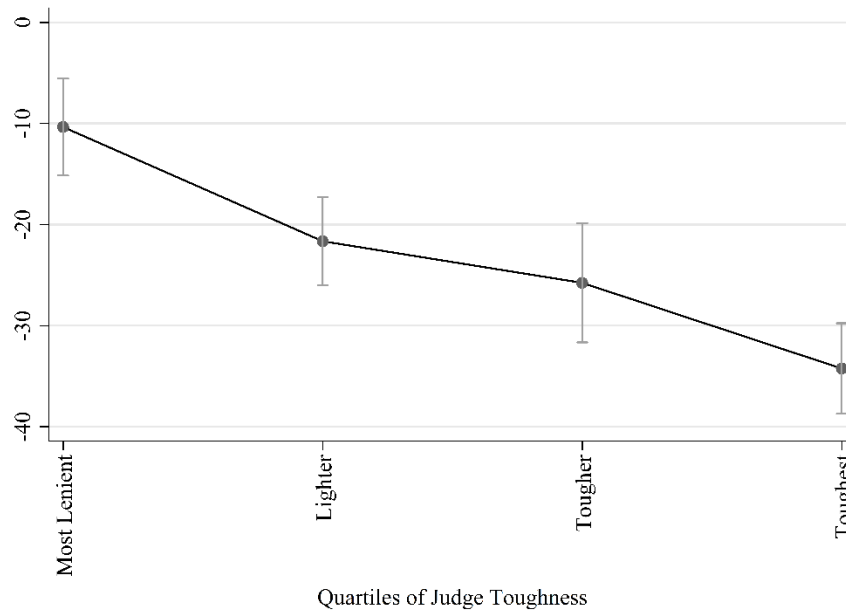
Thus, the dots for each quartile represent the percentage change in sentence length. The bars extending upwards and downwards from the dots represent a confidence interval. If the entire confidence bar is above or below zero, then we can say with some confidence that the imposition of voluntary sentencing guidelines exerted a statistically significant effect on the sentence length imposed for that quartile of judges. On the other hand, if the bar is both above and below zero, we cannot confidently say that the sentencing guidelines imposed a significant impact on decisions by that quartile of judges.

*Figure 17: Effect of Voluntary Sentencing Guidelines on Length of Sentences Imposed, by Quartiles of Judge “Toughness”*



We see that when the worksheet became voluntary, there was no significant change among the most lenient and the toughest judges. But the middle quartile of judges appears to have significantly reduced its sentence lengths after the imposition of voluntary sentencing guidelines. One explanation for this may be that judges on both extremes of the sentencing distribution are generally unwilling to change their behavior on a voluntary basis because of underlying personal beliefs, principles, or political pressures. By contrast, the most lenient judges may have already been so lenient in prior sentencing that the guidelines provided no additional latitude or motivation to further reduce sentence length. Whatever the explanation, the data suggest that voluntary sentencing guidelines have been most effective at changing the behavior of the middle two quartiles of judges. Figure 18 replicates this methodology for presumptive sentencing guidelines.

Figure 18: Effect of Presumptive Sentencing Guidelines on Length of Sentences Imposed, by Quartile of Judge “Toughness”



As seen in Figure 18, presumptive sentencing guidelines appear to have a statistically significant effect on all quartiles of all judges in Alabama. The result is a sort of gradient effect: the toughest judges prior to the policy change are required to dial back sentencing the most, and the most lenient judges, *ex ante*, move the least.

It is also important to evaluate the robustness of our findings. Since these policy changes became effective in October 2006 and 2013 around the end of the fiscal year,<sup>152</sup> there may be concern that other changes happened around the same time that may be influencing our results. It is worth reiterating that we know of no other changes that happened in either 2006 or 2013 that may affect sentencing decisions. Nevertheless, we recognize that there may be something about the change of a fiscal year driving the results. That is, there is some unobservable factor correlated with a fiscal-year change that is driving the results. To test for a “fiscal-year effect,” we alter equation (1) and run twelve distinct regressions to measure a placebo difference-in-differences estimate for each fiscal-year change between 2003 and 2014.<sup>153</sup> For instance, testing for a fiscal-year effect in 2009, the equation we estimate is:

152. See *supra* Figure 1 (showing the dates that each of these changes occurred).

153. Fiscal years 2002 and 2015 are omitted due to a lack of pre/post data.

$$O_{it} = a_1 \text{Group}V_{it} + a_2 \text{Post}V_{it} + \alpha \text{Group}V_{it} * \text{Post}V_{it} + b_1 \text{Group}P_{it} + b_2 \text{Post}2009_{it} + \beta \text{Group}P_{it} * \text{Post}2009_t + X_{it} + \varepsilon_{it} \quad (6)$$

where everything in equation (6) mirrors that of equation (1) except that the second group of policy variables that are enacted in 2013 are replaced with 2009 time and interaction variables. The results of these regressions are displayed graphically in Figure 19.<sup>154</sup> Note that these results are not measuring the year effect but rather a difference-in-differences effect as if the policy had changed in that fiscal year. So, for example, we would expect to see a significant result in 2006 and 2013 because those are the years when the policies actually changed. But, we would not expect to see a statistically significant effect in other years, as there was no policy change during those years. Like the previous figures, the lines that extend upwards and downwards from the dots for each year represent confidence intervals. Thus, if the bar is entirely above or below zero, we can say with some level of confidence that our result is significant—that is, we can feel confident that our results are not attributable to pure chance.

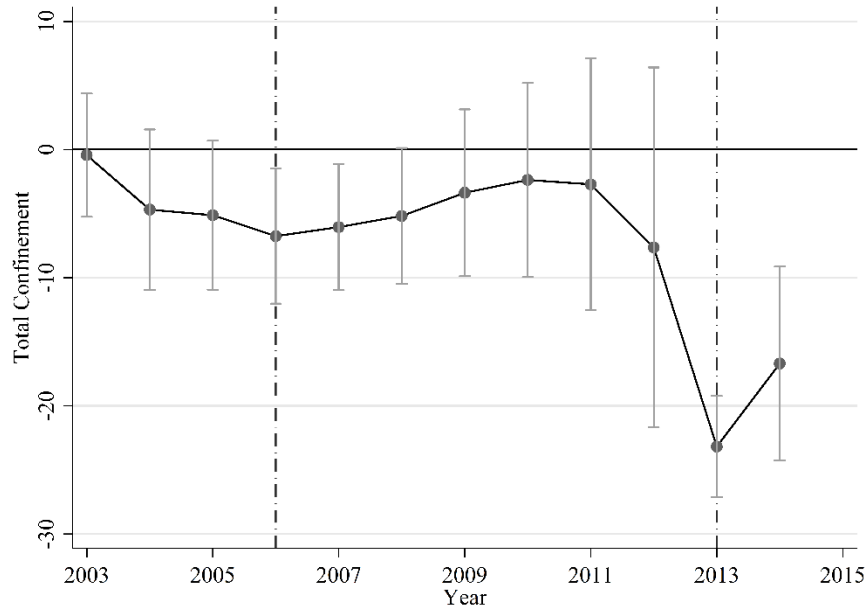
As seen in Figure 19, outside of the years where we expect to see a significant result, there is little evidence that the results are being driven by changing of the fiscal year.<sup>155</sup>

---

154. Corresponding tables are available upon request.

155. Since the difference-in-differences estimate essentially calculates a before-and-after average, we would expect to see results that near statistical significance as we approach the actual date of policy change, since the closer you get to the policy-change date, all that happens is the shifting of a couple data points from the pre- to the post-period.

Figure 19: Testing Placebo Year Effect



The nature of this data set and these policy changes provide us with an additional robustness check. While the worksheet applied to personal and burglary offenses on a voluntary basis starting in 2006, that classification of sentences did not become presumptive in 2013. Since we have a baseline of nonworksheet sentences, we are able to calculate a placebo difference-in-differences test on personal and burglary sentences in 2013 to ensure that the results we report are not capturing some unobserved change in the data.

To do this, we estimate a placebo difference-in-differences by restricting the data set to only include the nonworksheet sentences—to serve as the control group—and personal and burglary sentences. We then interact the personal-and-burglary dummy with the presumptive-post-treatment dummy. Again, since this class of sentences was not subject to the 2013 policy change, we would expect to see no effect. Replicating our earlier figures with this placebo test, we observe a statistically significant difference-in-differences estimate for only three of forty possible coefficients. This is about as frequently as we would expect to make a type I error at the 10% level.<sup>156</sup>

#### E. Limitations of Study

While we believe that our study provides compelling evidence about the usefulness of voluntary and presumptive sentencing guidelines, we recognize

156. The full tables of these results are on file with the authors and available upon request.



that it comes with some limitations. For one thing, Alabama is an imperfect case study because the state did not apply these sentencing guidelines to all criminal offenses. It is no surprise that a state like Alabama would be more inclined to experiment with presumptive sentencing guidelines for less serious drug and property offenses, but less inclined to make such reforms in cases of serious violent crimes like homicide.<sup>157</sup> It is hard to know how the type of offenses covered by these sentencing guidelines may influence the generalizability of our findings.<sup>158</sup>

It is also worth mentioning that our study does not fully consider the role of prosecutors in sentencing. In assessing the effect of the *Booker* decision on interjudge disparities, the U.S. Sentencing Commission observed that “[d]ifferences in charging and plea agreement practices at the district level have contributed to sentencing disparities.”<sup>159</sup> As an example, the Commission cited widespread variations in prosecutorial practices involving the filing of notices in drug-trafficking offenses, the charging of multiple violations under 18 U.S.C. § 924(c), and the use of binding plea agreements in order to recommend a particular sentence length.<sup>160</sup> Similarly, Professors Starr and Rehavi have made compelling arguments that some evidence of apparent sentencing disparity may result more from events that happen before sentencing, like prosecutorial decisions and plea-bargaining choices.<sup>161</sup> We recognize that our study cannot control for all of these potential variables. Nevertheless, we remain convinced that the changes we observe in our data are likely the result of mostly changes in judicial behavior, rather than changes in behavior by prosecutors. For one thing, we observe sudden and

---

157. To expand on this point, presumptive sentencing guidelines, by their very nature, may sacrifice individualization in favor of uniformity. To the extent policymakers believe that individualization of punishments is appropriate, we think it is likely that policymakers would most desire the ability of judges to individualize sentences to reflect moral culpability in the cases of the most serious, violent felony offenses. Additionally, there are other reasons why nonworksheets offenses exist and these reasons predate the rise of the voluntary-presumptive dichotomy—beginning with a tight development calendar for rolling out a guideline package, a limited workforce, and very little funding. All of this led to a focus on “high rate” crimes (i.e., the ones that we are having to address on a regular basis). Low-rate crimes and misdemeanors were shelved and left for another day, as were capital offenses, due to the complexity and the ever-changing rules for capital sentencing. Additionally, building on the approach of some others, the first round of guidelines tracked historical sentencing practices. High-rate crimes had a rich history from which to work; low-rate crimes did not. Using historical sentencing data, the Commission could better conform to the legislature’s challenging requirements for maintaining judicial discretion, reducing disparity, addressing prison overcrowding, and even reducing some prosecutorial discretion.

158. For example, we cannot say whether judges would respond the same were voluntary and presumptive sentencing guidelines enacted for all classes of criminal offenses. We also cannot discount the possibility that judges would be more likely to depart from the presumptive guidelines for serious criminal offenses like homicide, thereby diminishing some of the reductions in disparity we identified in this study. These limitations raise some questions as to the generalizability of our findings.

159. U.S. SENTENCING COMM’N, *supra* note 108, pt. A, at 7.

160. *Id.*

161. Starr & Rehavi, *supra* note 112, at 2.

obvious changes in sentence lengths almost immediately after the introduction of presumptive sentencing guidelines, as illustrated visually in Figure 3. We also saw sudden and obvious changes in disparities at this same time period, as illustrated visually in Figures 6 and 9. In addition, we have found no evidence of displacement. That is, we have found no evidence that prosecutors are simply charging criminal defendants with different offense categories to avoid being bound by the new guidelines.<sup>162</sup> Instead, all of the evidence in our comprehensive data set suggests that some classes of judges—particularly those with a history of issuing the toughest sentences—responded to the sentencing guidelines by issuing sentences more consistent with the sentencing of their peers for similarly situated defendants. And the data show that this has correlated with apparent reductions in the overall prison population and the percentage of the prison population serving sentences for nonviolent offenses.<sup>163</sup>

This study also does not consider the effect of a judge's race, gender, or political affiliation on sentencing behavior. Prior research strongly suggests that all of these factors may affect judicial sentencing decisions.<sup>164</sup> It may be helpful for future research to explore how these factors influence sentencing behavior of judges in Alabama and in other states under these voluntary and presumptive guidelines.

#### IV. Implications

Overall, we find that Alabama's implementation of voluntary and presumptive sentencing guidelines helped the state successfully drive down overall sentence length, reduce interjudge disparities, and reduce racial disparities in sentences. It is hard to see the Alabama Sentencing Commission's efforts as anything less than a success in achieving most of these ends. These findings have several implications for the literature on criminal sentencing. They suggest that, on balance, presumptive guidelines may be preferable to voluntary guidelines in regulating sentence lengths and reducing interjudge and racial disparities. To the extent that *Blakely v. Washington* created a "constitutional tax" on the use of presumptive

---

162. We show this in a number of ways—all of which are available upon request. First, we show that the distribution of indictment scores and ranks does not change in the weeks leading in and out of the policy changes. Second, in running the difference-in-differences regressions, we can show that the most serious indictment score and rank are not affected by the policy changes in any meaningful way. For instance, the sign on each policy change is negative in nearly every iteration of each regression. If prosecutors were trying to lobby for a longer sentence, the strategy would necessarily be to push for a harsher indictment that fell into the personal/burglary category that is only subject to voluntary guidelines or the nonworksheet category not subject to any guidelines—both of which carry higher indictment scores than drug/property crimes, so we'd expect to see the presumptive-guideline policy change lead to an increase in indictment score. This is not a phenomenon we observe in the data or the regressions.

163. For more information on these trends, see Part IV discussing these data in more detail.

164. See sources cited *supra* note 106.

sentencing-guideline systems,<sup>165</sup> our data suggest that the Court may have inadvertently contributed to more disparity in criminal sentences across the country. Finally, and perhaps most importantly, our data suggest that Alabama could serve as a blueprint for states looking to reform their sentencing laws. In an era when states are searching for ways to reduce disparities in criminal sentencing and face mounting economic pressures to reduce their prison populations, the Alabama model is particularly instructive. It shows how a state was able to use sentencing guidelines, particularly presumptive guidelines, to reduce disparities measurably and modestly cut its overall incarceration of nonviolent offenders.

A. *Importance of Presumptive Sentencing Guidelines*

To begin with, our findings suggest that voluntary guidelines alone may be somewhat effective in reducing sentence disparities. This is consistent with the prior work by other scholars like Professor Pfaff. Remember, Pfaff looked at prior attempts by a number of states to implement voluntary and presumptive sentencing guidelines.<sup>166</sup> He predicted that, while presumptive guidelines may reduce sentence variation by between 54% and 57%, voluntary guidelines would achieve much of this same reduction in variation—roughly between 21% and 35%.<sup>167</sup> Our study shows that, while voluntary guidelines may help reduce disparities, they remain comparatively less effective than presumptive guidelines. For one thing, we saw in Figures 15 and 17 that both voluntary and presumptive sentencing guidelines were associated with decreases in disparities between black and white offenders. Our models are highly confident that presumptive guidelines reduced racial disparities. Our models are somewhat less confident, though, that voluntary guidelines achieved this result—particularly when we added in controls. Our ability to track the behavior of individual judges gives us additional insight into how voluntary and presumptive guidelines may be influencing judicial behavior. As seen in Figures 18 and 19, voluntary guidelines appear to exert the greatest effect on the middle-two quartiles of judges in comparative punitiveness—that is, judges who have historically issued punishments around the median of their peers. By contrast, presumptive guidelines appear to be particularly effective at altering the behavior of judges who have historically been more punitive than their peers. Thus, we reach a relatively similar conclusion to Professor Pfaff: voluntary guidelines help, but presumptive guidelines are ultimately more effective at combatting inequality in our justice system.

---

165. *Blakely v. Washington*, 542 U.S. 296, 318 (2004) (O'Connor, J., dissenting).

166. Pfaff, *supra* note 7, at 255–85 (describing Pfaff's model and his results).

167. *Id.* at 235.

B. *The Costs of Blakely*

Our findings have particularly important implications because of the Court's decision in *Blakely v. Washington*.<sup>168</sup> Remember, in that case, the Court found that Washington's state sentencing guidelines violated the Sixth Amendment by forcing judges to make factual findings by a preponderance of evidence that increased the maximum permissible sentence range for a criminal offender.<sup>169</sup> At the time of *Blakely*, an estimated thirteen states had presumptive-sentencing-guideline systems that appeared to violate the Court's new understanding of the Sixth Amendment.<sup>170</sup> In the wake of this decision, commentators widely worried that states would simply abandon these presumptive sentencing guidelines rather than making them *Blakely*-compliant—which, as Justice O'Connor noted in her *Blakely* dissent, would force states to take on additional costs and burdens.<sup>171</sup> Alternatively, it seems possible that some states that would have adopted presumptive sentencing guidelines may have chosen not to do so after *Blakely* because of this added cost of complying with the Court's understanding of the Sixth Amendment. Today, according to one estimate, as few as five states and Alabama employ a truly mandatory- or presumptive-sentencing-guideline system.<sup>172</sup> This means that some states have likely abandoned their sentencing guidelines after *Blakely* or have moved to voluntary guideline systems.

To be clear, even after *Blakely*, states can still develop presumptive-sentencing-guideline systems that significantly reduce disparities and regulate sentence lengths. Obviously, Alabama has done just this. But to the extent that *Blakely* made the implementation of these presumptive-sentencing-guideline systems particularly taxing, time-consuming, or expensive, and discouraged their use, the decision may have inadvertently contributed to additional inequality in our criminal justice system. This is consistent with the literature on the impacts of *Booker* on criminal sentence lengths in the federal system. For example, much as Professor Yang found that *Booker* contributed to significant increases in sentencing disparities,<sup>173</sup>

---

168. *Blakely*, 542 U.S. at 303–04 (extending the logic of *Apprendi* to hold that state sentencing guidelines could not permissibly require trial judges to make factual findings on a standard lower than beyond a reasonable doubt and then use those factual findings to increase the maximum permissible sentence for a criminal offender).

169. *Id.*

170. Pfaff, *supra* note 7, at 250 (“While seventeen states (along with the federal government) employ some sort of determinate sentencing law or presumptive guideline system, *Blakely* affects only thirteen of them: Four states employ sentencing regimes that do not violate *Blakely*.”) (footnote omitted).

171. *Blakely*, 542 U.S. at 318–19 (O'Connor, J., dissenting) (arguing that the Court's decision in *Blakely* would “exact[] a substantial constitutional tax” on states by forcing them to conduct “full-blown jury trial[s] during the penalty phase proceeding” to determine an offender's sentence).

172. Mitchell, *supra* note 4, at 36 tbl.5 (showing a diagram of all states with sentencing guidelines and how these compare to one another).

173. Yang, *supra* note 1, at 1307.

our data suggest that some states may have disparities in their criminal-sentencing process that would not exist but for *Blakely*. Ultimately, though, we can only hypothesize on this point. Our data merely show the inverse: a state without sentencing guidelines (Alabama) was able to reduce disparities through the introduction of voluntary and presumptive guidelines. We believe it is likely that a state undergoing this process in reverse—that is, going from presumptive guidelines to voluntary or no guidelines—would experience an increase in disparities. And it seems likely that, to the extent any states opted against the use of presumptive sentencing guidelines because of *Blakely*, this may have contributed to more inequality in their justice systems. Nevertheless, more research is necessary to verify these hypotheses.

### C. *Alabama as a Blueprint*

Finally, based on our data, we believe that Alabama could serve as a blueprint for other states when it comes to the introduction of sentencing guidelines. According to an analysis of sentencing guidelines across the United States by Professor Kelly Lyn Mitchell, only a handful of states employ voluntary or presumptive sentencing guidelines,<sup>174</sup> and the overwhelming majority of American states do not employ sentencing guidelines.<sup>175</sup> It seems theoretically plausible that these jurisdictions could reduce interjudge and racial disparities in sentences through the use of sentencing guidelines similar to those employed by Alabama—particularly presumptive sentencing guidelines.<sup>176</sup>

Additionally, Alabama could also serve as a model for other states for another reason. Alabama was able to implement sentencing guidelines that substantially reduced interjudge and racial disparities, while also driving down overall sentence lengths. As the Alabama Sentencing Commission’s recent annual report implies, this has resulted in substantial cost savings for the state. The state has seen a “shift to a lower percentage of non-violent offenders in the State prison system,”<sup>177</sup> thereby freeing up limited resources for the state to use on the most high-risk and morally culpable offenders. Alabama will likely face increased pressure to further reduce prison populations in light of the recent federal investigation that found Alabama

---

174. Mitchell, *supra* note 4, at 36 tbl.5 (showing a summary of existing sentencing guidelines based on Professor Mitchell’s analysis). It is worth noting that Professor Mitchell found that Michigan and Pennsylvania have sentencing guidelines that “lean mandatory” but are still generally voluntary in nature. *Id.*

175. *Id.*

176. In fact, by using Professor Mitchell’s estimates, it seems that only around 33,027,876 individuals, or around 11% of the total U.S. population of 308,745,538 residents, live in states with some type of presumptive guidelines. This suggests that a substantial cross-section of the American population live in jurisdictions that could benefit from following the Alabama model.

177. ALA. SENTENCING COMM’N, 2017 REPORT vii (2017) [hereinafter 2017 REPORT], <http://sentencingcommission.alacourt.gov/media/1055/2017-annual-report.pdf> [https://perma.cc/TY34-VTHV].

prisons suffer from pervasive violence and inhumane conditions caused in part by “chronic overcrowding” and “severe understaffing.”<sup>178</sup> In response, Alabama in the coming years will likely face pressure to either reduce prison populations or build new facilities.

Alabama is likely not alone. The U.S. Supreme Court decision in *Brown v. Plata*<sup>179</sup> served as a wake-up call to states about the need to provide facilities capable of humanely housing convicted offenders.<sup>180</sup> In that case, a group of prisoners filed a class action lawsuit against the California prison system, alleging that the state was in violation of the Eighth Amendment ban on cruel and unusual punishment because of prison overcrowding.<sup>181</sup> At the time of the *Plata* decision, California housed around 156,000 inmates, “nearly double the number that California’s prisons were designed to hold.”<sup>182</sup> Because of this, California prisons failed to provide inmates in these overcrowded prisons with constitutionally acceptable medical- and mental-health care, resulting in “[n]eedless suffering and death.”<sup>183</sup> For example, as many as fifty sick inmates had been held together in a twelve-by-twenty-foot cage for up to five hours waiting on medical care in one particularly egregious incident.<sup>184</sup> And California had maintained these overcrowded levels since at least 2011, with as many as fifty-four prisoners forced to share a single toilet in some facilities.<sup>185</sup> The Court in *Plata* ultimately upheld<sup>186</sup> a lower court order to reduce the prison population substantially within two years.<sup>187</sup>

*Plata* was a signal to other states across the country, in part because of how other states compared to the apparent overcrowding in California. As the Alabama Sentencing Commission observed in 2012, Alabama at the time had a prison system “designed for less than 14,000” inmates, yet it housed over “25,400 inmates.”<sup>188</sup> This meant that Alabama prisons were “190% of

---

178. Debbie Elliott, *Justice Dept. Finds Violence in Alabama Prisons ‘Common, Cruel, Pervasive,’* NPR (Apr. 3, 2019, 2:29 PM), <https://www.npr.org/2019/04/03/709475746/doj-report-finds-violence-in-alabama-prisons-common-cruel-pervasive> [https://perma.cc/KB5J-AZFC].

179. 563 U.S. 493 (2011).

180. See, e.g., ALA. SENTENCING COMM’N, 2012 REPORT 11–14 (2012) [hereinafter 2012 REPORT], <http://sentencingcommission.alacourt.gov/media/1050/2012-annual-report.pdf> [https://perma.cc/W2CE-DDK6] (providing a detailed summary of the *Plata* case and the implications for the decision on Alabama prison policy).

181. *Plata*, 563 U.S. at 499–500. It is worth noting that, specifically, *Plata* came before the Court in order to answer a more particular, narrow question: “whether the remedial order issued by the three-judge court is consistent with requirements and procedures set forth in a congressional statute, the Prison Litigation Reform Act of 1995 (PLRA).” *Id.* at 500.

182. *Id.* at 501.

183. *Id.*

184. *Id.* at 504.

185. *Id.* at 502.

186. *Id.* at 545.

187. *Id.* at 509–10.

188. 2012 REPORT, *supra* note 180, at 11.

design capacity with 1 prisoner per 180 persons.”<sup>189</sup> By contrast, at the time of *Plata*, California was at “184% of capacity with 1 prisoner per 239 persons.”<sup>190</sup> In *Plata*, the Court ordered California to reduce the prison population to 137.5% of design capacity.<sup>191</sup> Policymakers in Alabama worried that, were a federal judge to similarly rule that Alabama’s prison system violated the Eighth Amendment, the state would need to release around 7,600 inmates.<sup>192</sup>

Within this legal environment, the Alabama sentencing guidelines have not only reduced interjudge and racial disparities in sentences, they have also helped respond to this important need for reductions in the overall prison population.<sup>193</sup> These guidelines may be contributing to the state’s slowly taking better control over its prison population. In 2013, around 28% of the in-house prison population was there for nonviolent offenses.<sup>194</sup> This fell to 26% in 2014,<sup>195</sup> 25% in 2015,<sup>196</sup> 24% in 2016,<sup>197</sup> and 22% in 2018.<sup>198</sup> And during this time period, the in-house prison population in the state declined from a high of 28,440 in 2003,<sup>199</sup> to 27,255 in 2005,<sup>200</sup> to 25,493 in 2013,<sup>201</sup> and to 20,185 in 2018.<sup>202</sup> Overall, between 2003 and 2018, Alabama saw its in-house prison population decline by around 29%. In the years before the implementation of sentencing guidelines, overcrowding forced the state to house as many as 1,485 inmates in private out-of-state prison facilities in

---

189. *Id.*

190. *Id.*

191. *Plata*, 563 U.S. at 501.

192. 2012 REPORT, *supra* note 180, at 11–12 (noting as well that a strict application of the holding from *Plata* to Alabama was unlikely, in part because *Plata* was premised on some unique deficiencies in the California system; but also noting that Alabama has a history of issues with prior prison-rights litigation).

193. *See supra* Figures 14, 15 (showing that voluntary and presumptive sentence lengths contributed to reductions in sentence lengths, with the effect being most significant for presumptive sentencing guidelines).

194. ALA. SENTENCING COMM’N, 2014 REPORT 16 fig.16 (2014) [hereinafter 2014 REPORT], <http://sentencingcommission.alacourt.gov/media/1052/2014-annual-report.pdf> [https://perma.cc/23GG-BZMP].

195. ALA. SENTENCING COMM’N, 2015 REPORT 18 fig.16 (2015) [hereinafter 2015 REPORT], <http://sentencingcommission.alacourt.gov/media/1053/2015-annual-report.pdf> [https://perma.cc/MHW5-QL4K].

196. ALA. SENTENCING COMM’N, 2016 REPORT 16 fig.16 (2016) [hereinafter 2016 REPORT], [https://sentencing.umn.edu/sites/sentencing.umn.edu/files/alabama\\_2016\\_final\\_report.pdf](https://sentencing.umn.edu/sites/sentencing.umn.edu/files/alabama_2016_final_report.pdf) [https://perma.cc/D8K4-ZYWL].

197. 2017 REPORT, *supra* note 177, at 16 fig.16.

198. ALA. SENTENCING COMM’N, 2019 REPORT 16 fig.16 (2019) [hereinafter 2019 REPORT], <http://sentencingcommission.alacourt.gov/media/1070/2019-annual-report.pdf> [https://perma.cc/N9XW-VTZL].

199. 2005 REPORT, *supra* note 125, at iii.

200. *Id.*

201. 2014 REPORT, *supra* note 194, at 16.

202. 2019 REPORT, *supra* note 198, at 16.

2003.<sup>203</sup> Overcrowding also forced the state to house as many as 1,448 state inmates in county jails as recently as 2005.<sup>204</sup> By significantly reducing its prison population, Alabama effectively eliminated these practices by 2018. Thus, it appears that Alabama may have used sentencing guidelines to reduce not only unwarranted disparity but also overall incarceration, particularly among nonviolent offenders.<sup>205</sup>

To be clear, the sentencing guidelines discussed in this Article did not solve the challenges that the state faced after *Plata* and the recent federal investigation. But by exerting downward pressure on the sentences of the “toughest” quartile of judges and decreasing overall sentence lengths, the sentencing guidelines seem to have helped Alabama better manage these challenges.

### Conclusion

Before Alabama began its experimentation with sentencing guidelines, it faced many of the same challenges as other states across the country. It had an ongoing problem with similarly situated offenders receiving disparate sentences.<sup>206</sup> It had overcrowded prisons<sup>207</sup> that could be the target of future structural reform litigation.<sup>208</sup> And these high levels of incarceration came at a significant cost to the state’s taxpayers.<sup>209</sup> The introduction of voluntary and presumptive sentencing guidelines in Alabama has not single-handedly solved any of these issues, but it has likely helped the state address many of these challenges more effectively. Our data suggest that Alabama has seen a decrease in interjudge and racial disparities, particularly after the implementation of presumptive sentencing guidelines.<sup>210</sup> The state has also seen reductions in overall sentence lengths, which have helped it reduce the percentage of the limited prison space allocated to nonviolent offenders.<sup>211</sup> Alabama’s experimentation with sentencing guidelines appears to be a

---

203. 2005 REPORT, *supra* note 125, at 10.

204. *Id.* at 9.

205. It is also important to acknowledge that in October 2018 through 2019, “Alabama prison inmates have not been able to be paroled early” because of a “moratorium” established by the Governor. Kelley Smith, *SPLC Concerned Early Parole Freeze Is Causing Prison Problems*, WHNT NEWS (July 3, 2019, 7:19 PM), <https://whnt.com/2019/07/03/splc-concerned-early-parole-freeze-is-causing-prison-problems/> [<https://perma.cc/J3VB-3NCL>]. This has caused a recent uptick in the prison population, which appears to be unrelated to the effects of the sentencing guidelines.

206. 2007 MANUAL, *supra* note 120, at 75–76 (describing the existence of “unwarranted sentencing disparities” at the time that Alabama implemented voluntary sentencing guidelines for the first time).

207. *Id.*

208. 2012 REPORT, *supra* note 180, at 11–14 (describing the *Plata* case and the threat of structural reform litigation).

209. 2007 MANUAL, *supra* note 120, at 75 (noting that overcrowding of prisons leads to “demands on our public resources”).

210. *See supra* subparts III(C)–(D).

211. *See supra* notes 193–202 and accompanying text.



success. But perhaps the most important lesson is not how the sentencing guidelines have improved the criminal justice system in Alabama, but what they tell us about sentencing guidelines more generally.

We believe our data set from Alabama allows us to make some tentative conclusions about the general usefulness of voluntary and presumptive sentencing guidelines at the state level. Our data provide evidence that, when it comes to reducing interjudge and racial disparities, and regulating sentence length, sentencing guidelines may be important.<sup>212</sup> Voluntary guidelines are useful, but presumptive guidelines seem more effective.<sup>213</sup> Presumptive guidelines appear particularly useful at influencing the behavior of outlier judges who have a history of issuing sentences substantially above or below their peers.<sup>214</sup> This realization underscores the potentially harmful consequences of *Blakely*.<sup>215</sup> To the extent that the decision has inhibited the expansion of presumptive sentencing guidelines across the country—either by discouraging some states from adopting presumptive guidelines because of their attendant costs, or by moving states from presumptive to voluntary sentencing guidelines<sup>216</sup>—the Court may have inadvertently contributed to more inequality in the American justice system.

---

212. *See supra* Part III.

213. *See supra* Figures 9–18.

214. *See supra* Figure 18 (showing that presumptive guidelines were comparatively more effective at changing the behavior of, a priori, the most punitive or lenient judges relative to their peers).

215. *See supra* subpart IV(C).

216. *See supra* subpart II(B) (describing the reasons why *Blakely* may have contributed to a reduction in the frequency of presumptive sentencing regimes).

Appendix: Summary Statistics

	Full Sample N=222,117		Non-Worksheet N=42,729		Personal & Burglary N=43,579		Drugs & Property N=135,809	
	Prior	Presumptive	Prior	Presumptive	Prior	Presumptive	Prior	Presumptive
Total Confinement	100.91 (127.98)	86.48 (106.37)	112.04 (147.08)	109.63 (112.05)	149.87 (217.61)	141.92 (194.21)	84.59 (74.17)	58.73 (68.95)
Confinement Imposed	48.14 (122.12)	37.46 (93.11)	53.03 (143.19)	42.73 (100.85)	86.22 (219.76)	77.14 (196.10)	36.32 (60.81)	22.97 (51.08)
Total Probation	26.70 (31.70)	26.86 (31.28)	29.11 (36.34)	28.94 (31.82)	26.20 (34.78)	26.11 (31.62)	26.28 (29.55)	25.92 (29.97)
Community Corrections	0.06 (0.23)	0.04 (0.20)	0.04 (0.19)	0.04 (0.17)	0.03 (0.18)	0.04 (0.19)	0.07 (0.25)	0.04 (0.20)
Most Serious Conviction Score	42.11 (24.81)	42.16 (26.43)	35.98 (36.09)	42.46 (38.54)	19.16 (18.86)	21.82 (19.59)	49.98 (17.62)	47.71 (19.81)
Most Serious Indictment Score	37.93 (25.54)	38.26 (27.77)	32.61 (35.25)	39.17 (38.54)	15.59 (18.19)	19.23 (19.51)	45.43 (20.06)	44.23 (21.55)
Plea	0.14 (0.35)	0.14 (0.34)	0.17 (0.38)	0.18 (0.37)	0.27 (0.44)	0.22 (0.42)	0.10 (0.30)	0.09 (0.29)
Indictment Counts	1.67 (2.28)	1.42 (1.63)	1.62 (1.63)	2.10 (34.25)	1.82 (1.74)	1.53 (2.77)	1.63 (2.53)	1.73 (1.98)
Number of Priors	0.91 (2.04)	1.17 (2.75)	0.82 (1.74)	1.43 (2.60)	0.96 (2.30)	1.17 (2.56)	0.92 (2.03)	1.48 (2.77)
Penitentiary	0.73 (0.44)	0.70 (0.46)	0.74 (0.44)	0.71 (0.45)	0.85 (0.36)	0.81 (0.39)	0.70 (0.46)	0.67 (0.47)
Split Sentence	0.29 (0.45)	0.28 (0.45)	0.31 (0.46)	0.33 (0.47)	0.36 (0.48)	0.36 (0.48)	0.26 (0.44)	0.22 (0.42)
Black	0.47 (0.50)	0.45 (0.50)	0.38 (0.49)	0.38 (0.48)	0.58 (0.49)	0.59 (0.49)	0.46 (0.50)	0.44 (0.50)
Male	0.82 (0.39)	0.81 (0.41)	0.84 (0.37)	0.80 (0.40)	0.91 (0.29)	0.90 (0.29)	0.78 (0.41)	0.76 (0.43)
Alcohol Anonymous	0.00 (0.06)	0.01 (0.08)	0.00 (0.04)	0.01 (0.09)	0.00 (0.03)	0.00 (0.03)	0.01 (0.07)	0.01 (0.08)
Drug Court	0.00 (0.03)	0.00 (0.04)	0.00 (0.02)	0.00 (0.03)	0.00 (0.00)	0.00 (0.02)	0.00 (0.03)	0.00 (0.05)
Mental Health	0.01 (0.07)	0.00 (0.07)	0.01 (0.07)	0.00 (0.07)	0.00 (0.06)	0.00 (0.06)	0.01 (0.07)	0.01 (0.07)
Anger Management	0.00 (0.07)	0.01 (0.12)	0.01 (0.09)	0.02 (0.13)	0.01 (0.12)	0.01 (0.10)	0.00 (0.03)	0.01 (0.10)
Habitual Offender	0.17 (0.37)	0.19 (0.31)	0.20 (0.37)	0.15 (0.36)	0.17 (0.38)	0.17 (0.38)	0.16 (0.37)	0.18 (0.39)
Drugs Near Housing	0.01 (0.09)	0.00 (0.06)	0.00 (0.05)	0.00 (0.03)	0.00 (0.00)	0.00 (0.01)	0.01 (0.11)	0.00 (0.08)
Drugs Near School	0.01 (0.10)	0.00 (0.07)	0.00 (0.05)	0.00 (0.03)	0.00 (0.01)	0.00 (0.01)	0.01 (0.12)	0.00 (0.02)