

Racial Disparities under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums

Joshua B. Fischman
*Max M. Schanzenbach**

* Associate Professor of Law, Northwestern University, joshua.fischman@law.northwestern.edu; Professor of Law, Northwestern University, m-schanzenbach@law.northwestern.edu. The authors thank Amy Baron-Evans, Rachel Harmon, Paul Hofer, Kate Stith, two anonymous referees, and participants at the Conference on Empirical Legal Studies, the American Law and Economics Association Annual Meeting, and the University of Chicago Law and Economics Workshop. An earlier version was presented at the Law and Economics of Race conference at the University of Chicago.

Abstract

The United States Sentencing Guidelines restrict judicial discretion in part to reduce unwarranted racial disparities. However, judicial discretion may also mitigate disparities if judges use discretion to offset disparities emanating from prosecutorial discretion or sentencing policies that have a disparate impact. To measure the impact of judicial discretion on racial disparities, we examine doctrinal changes that affected judges' discretion to depart from the Guidelines. We find that racial disparities are either reduced or little changed when the Guidelines are made less binding. Racial disparities increased after recent Supreme Court decisions declared the Guidelines to be advisory; however, we find that this increase is due primarily to the increased relevance of mandatory minimums, which have a disparate impact on minority offenders. Our findings suggest that judicial discretion does not contribute to, and may in fact mitigate, racial disparities in Guidelines sentencing.

I. Introduction

Twenty-five years ago, Congress approved the United States Sentencing Guidelines, imposing a highly structured sentencing regime on federal district courts and opening sentencing decisions for the first time to meaningful appellate review. One of the primary goals of the Guidelines was the reduction of unwarranted racial disparities in sentencing. Indeed, liberal stalwart Senator Ted Kennedy was a primary sponsor of the Sentencing Reform Act, and liberal interest groups testified in favor of the Act on the grounds that binding guidelines would reduce racial disparities (Stith & Koh 1993). Concerns about racial disparities in sentencing have resurfaced since the Supreme Court declared the Guidelines to be advisory and expanded judicial discretion in *United States v. Booker*¹ in 2005. Scholars and policymakers are debating whether binding guidelines should be reenacted in a modified form, and this debate has been influenced by concerns that racial disparities would increase under advisory guidelines.

Theory alone cannot determine how expanding judicial discretion affects racial disparities. Judges' decisions may be influenced by subconscious bias or racial stereotyping, in which case expanding judicial discretion would likely increase racial disparities. On the other hand, judicial discretion could reduce racial disparity by mitigating bias from other actors. For example, racial disparities could be the result of prosecutorial bias or sentencing policies that have a disparate impact against minorities.

To measure the impact of judicial discretion on racial disparities, we examine four changes in sentencing law over the last twenty years that affected judges' ability to depart from the Guidelines. We find that racial disparities were generally lower during periods when judges had wider discretion, suggesting that judges exercise discretion in a manner that mitigates disparity. This result is mostly confirmed when we examine shorter time windows around each doctrinal change; when judicial discretion is increased, racial disparities either decrease or do not change significantly.

In contrast to our main result, we find that racial disparities increased after the Supreme Court's 2007 decisions in *Rita v. United States*,² *Gall v. United States*,³ and *United States v. Kimbrough*⁴ (hereinafter "RGK"), which clarified the Court's holding in *Booker*

¹ 543 U.S. 220 (2005).

² 551 U.S. 338 (2007).

³ 552 U.S. 38 (2007).

⁴ 552 U.S. 85 (2007).

and further encouraged departures from the Guidelines. This finding cannot be attributed to judicial bias, however, because the rate at which mandatory minimums apply to minority offenders also increased after *RGK*. Thus, we conclude that judges were not able to reduce sentences for minority offenders as much as they could for whites.

To partially isolate the effect of mandatory minimums, we divide the cases into two categories: those in which the minimums were unlikely to be binding (offenders with low criminal history scores and offenses not involving firearms) and those in which minimums were likely to be binding (offenders with substantial criminal histories and cases involving firearms). The increase in racial disparities after *RGK* occurred primarily in those cases in which mandatory minimums are more likely to be binding. Thus, we find no evidence that the increase in racial disparities after *RGK* is due to the biased exercise of judicial discretion.

Prior studies examining the impact of *Booker* on racial disparities generated conflicting conclusions. The United States Sentencing Commission (2006, 2010) conducted two studies that found that racial disparities increased after the Guidelines were declared advisory. But Ulmer et al. (2011) applied an alternative empirical model that found no significant increase in racial disparity after *Booker*. One important difference between our findings and the prior results is that the prior studies did not directly address the contribution of mandatory minimums to racial disparities. A second difference is that we only include explanatory variables that are exogenous to the sentencing judge's decision. The Commission and Ulmer et al. controlled for judge-induced departures, whether a mandatory minimum applied, and the presumptive guidelines minimum. These are not control variables, but rather outcome variables, which were themselves influenced by *Booker* and *RGK*.

Taken together, our results support two important conclusions. First, racial disparities under the Guidelines are not attributable to judicial discretion; in fact, judicial discretion likely reduces racial disparities, at least in the current federal Guidelines regime. Second, mandatory minimums have become more important in the post-*Booker* era and may need serious reevaluation in light of the potential disparities they produce.

The remainder of this article is organized as follows. Part II explains the Guidelines and the legal changes in the standard of review, Part III explains our empirical strategy and describes the data, Part IV discusses the results, and Part V concludes.

II. The Guidelines, Mandatory Minimums, and Standards of Review

The Sentencing Reform Act of 1984 established the United States Sentencing Commission and charged it with developing guidelines that would “provide certainty and fairness in meeting the purposes of sentencing, avoiding unwarranted disparities among defendants with similar records who have been found guilty of similar criminal conduct.”⁵ Pursuant to this directive, the Sentencing Commission developed a sentencing table, displayed in Appendix Figure 1.

The sentencing calculation begins with a defendant’s base offense level, which is determined by the crime of conviction. A district judge may then make adjustments to this offense level on the basis of factors specified within the Guidelines—such as whether the offender played a major or minor role in the crime—resulting in a final offense level. The offender’s criminal history category is determined by the offender’s prior offenses. Together, the offense level and criminal history yield a sentencing range. At higher ranges, the lower bound of the sentencing range is 75 percent of the upper bound. Prior to *Booker*, any properly calculated sentence within the Guidelines range was unreviewable, except on very narrow grounds.⁶ Calculation of the sentencing range itself was reviewable, but mostly under a “clear error” standard because of the fact-intensive nature of offense level calculations.

A judge may also depart from the guidelines if he or she finds that the case presents atypical circumstances that were not contemplated by the Guidelines. The judge must give reasons to justify this decision, and departures may be appealed by the prosecutor or the defendant. Since the enactment of the Guidelines, departures have been reviewed *de novo* and deferentially at different points in time.

Over 90 percent of convictions are the product of plea bargains, but judges nevertheless exert a significant influence over offenders’ final sentences. Plea bargains are made in the shadow of the sentencing judge, whose identity will be known at the time of the plea. In addition, sentencing judges are not bound by factual stipulations in plea agreements; the judge may make adjustments to the plea-bargained base offense level on the basis of additional findings of fact. Finally, the judge may grant a departure if warranted by the circumstances of the case. LaCasse and Payne (1999) found empirical evidence that judges exerted a strong influence over the substance of a plea agreement, and that this influence persisted post-Guidelines. Likewise, Fischman and Schanzenbach (2011) and Schanzenbach and Tiller (2007) found that judicial

⁵ 28 U.S.C. § 991(b).

⁶ For example, if a judge stated that he did not think a departure was permitted on the grounds offered by the defense, an appellate court could remand with instructions to consider a departure.

characteristics affect the calculation of the offense level, the probability of a departure, and the overall prison sentence.

Even though judges still exert a substantial influence on sentencing under the Guidelines, most studies conclude that the Guidelines reduced inter-judge sentencing differences. Two prominent studies found that the introduction of the Guidelines reduced inter-judge disparity (Anderson, Kling, and Stith 1999; Hofer, Blackwell, and Ruback 1999). Fischman and Schanzenbach (2011) found that inter-judge disparity under the Guidelines depended on the standard of appellate review for sentencing departures: differences between Democratic and Republican appointees were significantly larger under deferential review than under de novo review. Scott's (2010) study of a single district found that inter-judge disparities increased after the Guidelines were declared advisory in *Booker*. It does not follow, however, that increases in inter-judge disparity correspond to increases in racial disparity.

Despite the Guidelines' reduction of inter-judge disparity, most studies have concluded that large racial disparities persist under the Guidelines. McDonald and Carlson (1994) found that racial disparities in sentencing increased after the adoption of the Guidelines, but they attributed most of this effect to the difference in mandatory minimums for crack and powder cocaine. In a detailed study of post-Guidelines sentencing, Mustard (2001) found substantial unexplained racial differences even after controlling for detailed offense characteristics, with blacks receiving one-third fewer departures than whites and 10 percent longer prison sentences than whites. What causes the observed disparities has not been convincingly established, but there is some reason to be skeptical that judicial temperament is an important factor. Using district-level variation, Schanzenbach (2005) confirmed the existence of racial disparities in sentencing but found no evidence that these disparities were affected by the race, gender, or political orientation of the judges in a district.⁷

A. Standards of Review for Departures

Since the adoption of the Guidelines, there have been several changes to the standard of review for sentencing departures. The Sentencing Reform Act initially directed courts of appeals to "give due deference to the district court's application of the guidelines to the facts" (18 U.S.C. § 3742(e)(4)). In the early years of the Guidelines, departure decisions were typically treated as mixed questions of law and fact, with a different standard of review applied to each determination. Factual determinations that formed the basis for the departure were reviewed for clear error. Whether the facts were

⁷ There are some studies assessing the impact of the PROTECT Act and *Booker*, though not the effect on racial disparities. Freeborn and Hartman (2010) find no evidence that the PROTECT Act changed sentencing practices apart from reducing departure rates. Scott (2010) concluded that there was an increase in inter-judge disparity in one district after *Booker*.

sufficiently unusual to justify a departure from the Guidelines was reviewed de novo. Finally, the magnitude of a departure was reviewed for abuse of discretion. Most circuits followed this approach until 1996.⁸

In *Koon v. United States*,⁹ the Supreme Court adopted an abuse-of-discretion standard for review of departures. This standard remained in place until Congress enacted the PROTECT Act of 2003,¹⁰ which reinstated a de novo standard of review for the application of the facts to the Guidelines in departure cases. The Commission responded to the PROTECT Act by imposing a moratorium on new grounds for departures and imposed some reporting requirements on prosecutors to encourage appeals of downward departures.

In 2005, the Supreme Court held in *Booker* that the Guidelines could not be binding. In order to satisfy the Sixth Amendment's guarantee of trial by jury, the Court held that sentencing judges must take the Guidelines to be merely "advisory." The Guidelines retained some degree of force, however, since the Court preserved appellate review of sentences under a "reasonableness" standard.

In 2007, the Supreme Court further clarified *Booker* in a series of three cases. Although *Booker* had declared sentences within the Guidelines range to be reviewable, *Rita* held that courts of appeals may attach a presumption of reasonableness to these sentences because they reflect the judgment and expertise of both the sentencing judge and the Sentencing Commission. *Gall* held that courts of appeals must review all sentences—those outside the Guidelines range as well as those within it—under a deferential abuse-of-discretion standard. However, *Gall* reaffirmed that reviewing courts could use the Guidelines as their starting point, and that the greater the departure from the Guideline range, the more searching the review could be. In a more dramatic shift, the Court held in *Kimbrough* that a district court could depart from the Guidelines if it disagreed with the policy choices embedded within them—in this case the disparity in treatment between crack and powder cocaine.

⁸ One exception for the First Circuit, which adopted a "reasonableness" standard in 1993. See *United States v. Rivera*, 994 F.2d 942, 951–52 (1993). The tripartite standard of review appears to have become less consistent in other circuits after *Rivera*. See, e.g., *United States v. Canoy*, 38 F.3d 893, 908 (7th Cir. 1994) ("when a district court clearly explains the basis for its finding...[it] is entitled to considerable respect on appeal") (citing *Rivera*); *United States v. Simpson*, 7 F.3d 813, 820 (8th Cir. 1993) (favorably citing *Rivera*).

⁹ 518 U.S. 81 (1996).

¹⁰ Prosecutorial Remedies and Other Tools to End the Exploitation of Children Today Act of 2003 (PROTECT Act), Pub.L. No. 108-21, (codified and amended at U.S.C., tit. 18 § 2423 (a)-(g), 2246, 2516(1)(c), 1591, 3142(e), 3283, 3559(e)), section 401.

Booker, Rita, Gall, and Kimbrough substantially weakened the force of the Guidelines, but they left the mandatory minimum sentences intact. By encouraging judges to sentence below the Guideline range, these decisions increased the importance of the mandatory minimum penalties. Prior to *Booker*, roughly 6 percent of offenders were sentenced at the mandatory minimum. By 2009, this level had increased to nearly 11 percent.

B. Mandatory Minimums

Shortly before the Guidelines went into effect, Congress expanded mandatory minimum sentences for drug and firearms offenses.¹¹ These minimums typically mandate sentences of at least 5, 10, 15, and 20 years, depending on the circumstances of the offense. The most controversial of the mandatory minimums are for drug crimes, because of their severity and the striking disparity of treatment between crack and powder cocaine: each gram of crack is considered equivalent to 100 grams of powder cocaine in calculating minimum sentences.

The Guidelines sentences were calibrated by the Sentencing Commission to be higher than statutory minimum penalties (United States Sentencing Commission 2011). The statutory minimums thus place a hard limit on the extent to which judges can reduce sentences through adjustments to offense levels or departures from the Guidelines. Nevertheless, these minimums often permit discretion in their application. For example, drug quantity is often an issue of serious contention in large conspiracies or when trafficking has occurred over time. In these circumstances, the mandatory minimum that should apply may be contested by the parties and potentially subject to negotiation. The prosecution exercises considerable discretion over whether to attempt to prove or charge quantities that would invoke the mandatory minimum.

Because Guidelines sentences were calibrated to be consistent with the minimum penalties, they incorporate the same crack-powder disparities. The Guidelines provided some flexibility by allowing judges to adjust or depart downward in appropriate circumstances, but not below the applicable mandatory minimum. In 2007, the Commission amended the Guidelines to reduce the weighting of crack relative to powdered cocaine to 20 to 1, and Congress followed suit in 2010 regarding mandatory minimums, reducing the ratio to 18 to 1.¹²

Apart from contested factual issues, there are three ways in which offenders may avoid minimum penalties. First, prosecutors may decline to charge defendants with

¹¹ Anti-Drug Abuse Act of 1986 (Pub. L. 99-570, 100 Stat. 3207); Pub. L. No. 98-473, § 1005(a), 98 Stat. 1837, 2138-39 (1986)

¹² Fair Sentencing Act of 2010 (Pub. L. 110-220).

offenses that are indicated by their conduct. Often, a defendant will plead to a lesser set of charges so as to avoid the imposition of a mandatory minimum. An early report by the United States Sentencing Commission (1991) found that 26 percent of defendants who were eligible for a mandatory minimum were not charged for the conduct that would subject them to the highest applicable minimum penalty. Schulhofer and Nagel (1997) subsequently estimated that minimums were circumvented in 20–35 percent of cases, although this varied substantially by district.¹³

Second, prosecutors can request “substantial assistance” departures below the mandatory minimum when the offender serves as a cooperating witness. Bowman and Heise (2001) show that the rates of substantial assistance departures vary dramatically across districts, and argue that these departures can be used to circumvent the Guidelines and mandatory minimums in addition to eliciting cooperation.

Finally, in certain drug cases, judges have some control over the application of mandatory minimums. Congress provided the “safety valve” departure in 1994, which grants judges the discretion to sentence offenders below the mandatory minimum if, among other conditions, the offender did not employ violence and does not have a criminal history score greater than one.¹⁴ Judges sometimes extended the safety-valve option to offenders with higher criminal history scores by making creative use of Guidelines-based departures for overstated criminal history. The PROTECT Act, however, forbade judges from using findings of overstated criminal history in the context of safety valve departures.¹⁵

The impact of mandatory minimums varies by criminal history score and offense type. In drug cases, mandatory minimums are less relevant for offenders with low criminal history scores, since these offenders are eligible for a safety-valve departure at the instigation of the judge. For drug offenders with lengthier criminal histories, mandatory minimums are more likely binding. Such offenders may also be subject to the “armed career offender” minimum,¹⁶ which specifies a 15-year sentence for those who have committed crimes employing firearms in the past. The prosecution, however, can decline to seek this minimum. Firearms offenses also have a variety of associated mandatory minimum penalties (for a description, see United States Sentencing Commission 2011: 269–93). Mandatory minimums are often not at issue or less binding for offenses not involving drugs or firearms.

¹³ Bjerk (2005) also finds evidence that prosecutors manipulate counts to avoid mandatory minimum penalties in state courts.

¹⁴ 18 U.S.C. §3553(f).

¹⁵ PROTECT Act § 4A1.3(b)(2)(B), (3)(A) and (3)(B).

¹⁶ 18 U.S.C. § 924(e).

One effect of mandatory minimums has been to transfer discretion from judges to prosecutors. Judges can unilaterally avoid the minimum penalties only when an offender is eligible for a safety-valve departure. Charge bargaining and substantial assistance departures, on the other hand, can be initiated only at the discretion of the prosecution, although the identity of the sentencing judge may influence the process. Unlike judges, prosecutors are not required to give reasons when they exercise discretion, and their decisions are not subject to significant appellate review.

Some commentators have suggested that prosecutorial discretion to enforce minimum penalties is exercised in a manner that exacerbates racial disparities. The U.S. Sentencing Commission (1991) qualitatively categorized cases in which a mandatory minimum could have been applied, and found that 46 percent of white defendants were sentenced below the mandatory minimum, compared to 32 percent of black defendants. Similarly, Meierhofer (1992) reported that relative to white offenders, black offenders were 21 percent more likely to be sentenced at or above the minimum, and Hispanics were 28 percent more likely. Maxfield and Kramer (1998) also found that white offenders were significantly more likely to receive substantial assistance departures than minority offenders, and that white offenders also received slightly larger sentence reductions for substantial assistance.

C. *Previous Estimates of the Impact of Booker*

The United States Sentencing Commission has produced two reports (2006, 2010), both asserting that *Booker* led to an increase in racial disparities. This conclusion was disputed by Ulmer et al. (2011), who find no significant impact of the PROTECT Act or *Booker* on sentencing practices.

We believe that the Commission's reports and the study by Ulmer et al. (2011) both have methodological problems. The primary flaw is that these studies include controls – such as presumptive Guidelines sentence, whether there was a judge-induced departure, or whether a mandatory minimum was binding – that are influenced by the judge's exercise of discretion or the policy change itself. These variables are endogenous, because they are causally influenced by judicial discretion and the sentencing reforms. We analyze these variables as outcome variables, not as controls.

The most important effects of *Booker* and *RGK* are that they increase judges' discretion to depart below the Guidelines range, and in fact, departures increased dramatically after *Booker*. By including controls that incorporate the effects of judge-induced departures, these regressions are only measuring the impact of *post-departure* discretion on racial disparity. Thus, they fail to capture the impact of the primary channel by which *Booker* and *RGK* affect sentences.

Similar critiques apply to some other variables included as control variables by the Sentencing Commission (2006, 2010) and Ulmer et al. (2011), such as final offense levels, presumptive sentence, and the mandatory minimum penalty. Offense level calculations are often fact-intensive and to some extent, subject to judicial and prosecutorial discretion (see Schanzenbach and Tiller 2007 for a discussion). Final offense levels increased after *Booker*, possibly because judges began to use departures instead of adjustments to reach a target sentence, or because prosecutors sought higher offense levels to counteract judges' greater discretion to depart.

Mandatory minimums are similarly endogenous to other sentencing outcomes. When the Guidelines become less binding, judges will be more likely to sentence at the mandatory minimum for two reasons. First, prosecutors have greater incentives to invoke mandatory minimums in order to retain greater control over the sentence. Second, judges who are now freer to sentence offenders outside the Guidelines range will be more likely to be constrained by the statutory minimum.

A prominent point of disagreement between the Commission's (2006, 2011) reports and Ulmer et al. (2011) involved whether to include offenders who received a sanction other than incarceration. The Commission included all such offenders, counting their sentence length as zero, while Ulmer et al. dropped these offenders from their analysis. The approach taken by Ulmer et al., which is consistent with many studies in the criminology literature (see studies cited in Ulmer et al. 2011), is motivated by a belief that decisions regarding incarceration are conceptually distinct from decisions involving sentence length.¹⁷

We follow the Commission's approach. For the purpose of measuring racial disparity, the population of interest consists of all convicted offenders, not merely those subject to incarceration. Sentencing reforms that impact the racial disparity in incarceration rates are just as relevant for normative purposes as reforms that impact the disparity in average sentence. By excluding non-incarcerated offenders, the approach of Ulmer et al. is introducing selection bias where it had not previously existed.¹⁸

¹⁷ As Bushway & Piehl (2001) point out, the convention of separately modeling incarceration and sentence length decisions originated before the widespread adoption of sentencing guidelines. The two-stage approach is less appropriate in a structured sentencing regime in which a judge must first calculate a sentencing range and then select a prison term within that range.

¹⁸ Ulmer et al. (2011) also provide alternative estimates from a tobit model that combines the incarceration and sentence length decision. The tobit model, however, assumes that the observed sentence length is derived from a latent variable that is censored at zero, and Ulmer et al. do not provide a theoretical basis for this latent variable. We agree with the Commission's use of ordinary least squares regression, which provides valid estimates of the average impact of explanatory variables on sentence length.

III. Empirical Estimation of Racial Disparities

There is little doubt that large unexplained racial disparities persist under the Guidelines, but the root causes of these disparities are not well understood. Some scholars have emphasized subconscious bias or racial stereotyping on the part of judges (Albonetti 1997; Steffensmeier et al. 1998; Everett & Wojtkiewicz 2002; Etienne 2006; Wang et al. 2009), while others have attributed disparities to prosecutorial bias (Maxfield and Kramer 1998; Lu 2007) or sentencing policies that have a disparate impact against minorities (Scott 2011; Baron-Evans & Stith 2012; Hofer 2012; Tonry 2012). Disparities stemming from judicial or prosecutorial bias would clearly be unwarranted; whether other sources of disparity are warranted may be contestable. Nevertheless, many observers of the criminal justice system (e.g., Roberts 2004; Provine 2007; Alexander 2010; Tonry 2012) believe that the current black-white sentencing gap is largely unwarranted, especially for drug crimes.

A. Empirical Strategy

In our empirical approach, we use the term “racial disparity” to describe the difference in sentencing outcomes between whites and non-whites after controlling for other observable offense characteristics. We then examine how racial disparity changes as doctrinal changes enhanced or limited judicial discretion.

We focus on the change in disparity for two reasons. First, it sheds light on the root causes of racial disparities, a subject of substantial controversy. If racial disparities do not change much when judges have greater discretion, then judicial bias is not likely to be the primary source of the disparities. Second, the estimated racial disparity is, by itself, difficult to interpret. Empirical study cannot distinguish between warranted and unwarranted disparity. Many studies of sentencing disparity control for myriad factors considered by the researcher to be legitimate and then presume that the racial effects remaining constitute unwarranted disparity. This approach relies upon several untenable assumptions. First, it assumes that there are no unobserved legitimate sentencing factors that correlate with race. Second, the approach assumes that none of the included control variables are themselves tainted by racial discrimination. For example, offense severity calculations could be tainted by prosecutorial or judicial bias, and black offenders’ criminal history could be overstated due to discriminatory policing and charging practices. Third, the included controls must provide normative justification for the differences in sentencing outcomes. For instance, although courts have consistently upheld the disparate treatment of crack and powder cocaine (see Sklansky 1995 for a discussion), the Sentencing Commission (2007) and many other commentators have argued that this disparity is unwarranted (e.g., Sklansky 1995; Roberts 2004; Provine 2007; Tonry 2012).

We examine racial disparities in several key outcome variables: downward departures, log prison sentence, final offense level, and whether an offender was sentenced at the mandatory minimum. By “downward departure,” we mean judge-induced departures, and therefore code substantial assistance departures as non-departures. Prison sentence is the most important substantive outcome; in addition we separately examine racial disparities in departures, offense levels, and the probability of receiving a sentence at the mandatory minimum. We take the log of prison sentence because the actual distribution of prison sentences, described below in greater detail, follows roughly a lognormal distribution. This is a result of the Guidelines’ prescription of prison sentences that are roughly exponential functions of the offense level calculation. We therefore can interpret our prison sentence results in percentage terms.

Our linear regression equation takes the following form:

$$Outcome_{ijt} = a + \beta_1 Black_{ijt} + \beta_2 Hispanic_{ijt} + \delta_1 Deferential_t \times Black_{ijt} + \delta_2 Deferential_t \times Hispanic_{ijt} + \lambda OffenderDemog_{ijt} + \mu OffenseType_{ijt} + \nu Grid_{ijt} + \rho District_j + \varepsilon_{ijt}$$

where i indexes offender, j indexes district, and t indexes the year and month of sentencing. *Outcome* represents our variable of interest. In the case of departures, *Outcome* equals one when a judge chooses to depart downward from the Guidelines, and zero otherwise. In the case of binding minimums, *Outcome* equals one if the sentence is at the statutory minimum. *Black* and *Hispanic* are dummy variables for black and Hispanic offenders. *Deferential* equals one when an “abuse of discretion” or “reasonableness” standard applies to departure decisions (between *Koon* and the PROTECT Act and after *Booker*) and zero when review is de novo. *OffenderDemog* includes offender demographic variables (age, age-squared, sex, citizenship status, and dummies for one, two, and three or more dependents). *OffenseType* includes dummies for offense category (divided into fifteen separate categories). *Grid* includes controls for criminal history, base offense level, and the interaction between the two, and *District* is a vector of dummies for each United States District Court. The main coefficients of interest are δ_1 and δ_2 , which are the interactions between the race variables and deferential review.

It is possible that secular trends in sentencing or changes in offender characteristics could affect our study, which encompasses a twenty-year time frame. To address time trends in our initial specification, we include a full set of interaction terms between *District*, *OffenseType*, and dummies for Guidelines year and *Deferential*. These time-district-offense interactions absorb a great deal of variation that may affect sentences, such as the changing nature of the underlying severity of a crime category and changes in district-level characteristics, including the composition of the bench.

We also perform separate regressions that examine a three-year window around each reform. Since our initial regression examines 17 years of Guidelines sentencing, it is possible that secular trends in sentencing could confound the effects of deferential review. Analyzing narrow time frames allows us to isolate the impact immediately around each doctrinal change and to determine whether these changes had different impacts. Changes to the Guidelines could change also prosecutorial strategies over time and the underlying mix of cases selected by prosecutors. By examining a three-year window around each reform, we control for these selection effects.¹⁹

The regression specifications employed here are reduced-form regressions of the judge's sentencing decision. We examine the logarithm of the total prison sentence,²⁰ downward departure, final offense level, and whether there was a binding minimum as outcome variables. We control for offense category and base offense level, which are determined by the crime of conviction and are therefore exogenous to the judge's decision. Because base offense levels incorporate the Guidelines penalties for crack and powder cocaine, including these controls means that the coefficients on the *Black* and *Hispanic* variables will not reflect the disparate impact of the crack-powder disparity on minority offenders.

Estimating the effects of mandatory minimum penalties is challenging because the presence of a binding minimum is endogenous. To isolate the effects of mandatory minimums, we divide the cases into two categories according to exogenous characteristics that predict whether minimum penalties are likely to be binding. In the "less binding" category are offenders in the lowest criminal history category who did not use firearms. The "more binding" category consists of drug offenders with higher criminal history scores and offenders who used firearms. If mandatory minimums are a dominant factor in explaining racial disparities, then we should observe larger increases in racial disparities in those cases for which mandatory minimums are more likely to be binding.

B. Data

We use Guidelines data from 1992 to 2009, the most recent year available.²¹ Since all convicted federal offenders must be sentenced, there is no concern that selection bias

¹⁹ The median time from indictment to sentencing is about six months, although there is heterogeneity by district. Moreover, many of these cases, particularly drug and firearm trafficking, are the products of preceding investigations and sting operations. Thus, it is unlikely that the change in the law itself affected the composition of cases brought in the narrow period examined.

²⁰ We add one to the prison sentence, measured in months, to avoid taking the logarithm of zero.

²¹ The Guidelines years prior to 1991 are excluded because of uncertainty about the standard of review, coding issues, and the presence of substantial numbers of non-Guidelines cases.

in litigated cases could confound our estimates.²² We limit our analysis to more serious crime categories, which have substantially higher sentences. These crime categories—murder, manslaughter, sex crimes, robbery (primarily bank robbery), drug trafficking, firearms offenses, racketeering, arson, and auto theft—constitute the 55 percent of all prison sentences but over 83 percent of all prison time imposed under federal criminal law. Moreover, over 94 percent of convictions for these serious crimes result in incarceration.²³ Less serious crimes, such as fraud and larceny, often start with low base offense levels resulting in Guidelines sentencing ranges that do not require a prison sentence. There is no reason for a judge to depart in such a case, since the range prescribed by the Guidelines already includes a no-prison option. Moreover, in these crime categories there are commonly sentencing alternatives available, such as home confinement. Thus, doctrinal changes that affect judges’ authority to depart from the Guidelines would be unlikely to have an impact in such cases.

Two further data issues remain. First, substantial assistance departures are initiated by the prosecution and must involve cases in which cooperation is of value, such as conspiracies. The role of judicial discretion in these cases is strongly curtailed. Recall that we use the term “departure” solely to refer to judge-induced departures.²⁴

Second, “fast-track” departures, granted pursuant to deportation, were especially common in the five southern border districts of Arizona, southern California, New Mexico, and south and west Texas. Moreover, the southern border districts are atypical due to large caseload pressures. Fast-track departures were codified by the PROTECT Act but are still of varying relevance across districts, and thus may confound the estimation.

We report results from regressions that include border districts as well as cases with substantial assistance departures, since we are most concerned with the average change in racial disparity. In unreported specifications, we excluded border districts and substantial assistance departures and found similar results.

²² Acquittals could conceivably create selection effects, however, these account for roughly one percent of federal criminal cases.

²³ For this reason, the disagreement between the Sentencing Commission (2006, 2010) and Ulmer et al. (2011) over whether to include non-incarcerated defendants is less central to our analysis. Nevertheless, we firmly believe that the Commissions’s approach of including all sentences is correct.

²⁴ A significant change, such as when the Supreme Court declared the Guidelines to be advisory in *Booker*, could lead prosecutors to offer more substantial assistance departures and hence substantial assistance departures may be endogenous to the legal change. In an earlier article, we found no evidence that the substantial assistance departure rate responds to changes in doctrine (Fischman & Schanzenbach 2011).

Appendix Table 1 details the summary statistics. Serious crimes are dominated by drug trafficking (70 percent), bank robbery (5 percent), and firearms offenses (16 percent). There is very little interpersonal crime, so victim characteristics are not a relevant factor in explaining racial disparities. Almost half of serious offenders have significant criminal histories (criminal history category greater than one). Mandatory minimums will be especially relevant for these offenders. Finally, 33 percent of offenders are white, 23 percent are Hispanic, and 42 percent are black. There are more than 500,000 sentences for serious crimes in the data set, which provides statistical power to detect even small differences in racial sentencing patterns.

IV. Results

We first present graphical evidence of the effect of the doctrinal changes, and then turn to our baseline regressions. Because the graphs suggest that there are long-run trends in sentencing, we next examine three-year windows around the changes in standards of review to test the short-run response of sentencing practices. In each case, we estimate the change in the black-white or Hispanic-white disparity in departure rates, prison sentences, offense levels, and proportion sentenced at the statutory minimum.

[FIGURE 1 ABOUT HERE]

A. Sentencing over time

Figures 1–3 depict departure rates, prison sentences, and binding minimums over the sample time frame for blacks and non-Hispanic whites.²⁵ Figure 1 shows the raw departure rate among serious crimes for blacks and whites. The departure rate for blacks is generally two to three percentage points below that of whites. There is a contraction in the black-white gap immediately after *Koon*, but the disparity opens up again after a few years. There is a decrease in the departure rate for both groups during the PROTECT Act, and a dramatic increase after *Booker* and another increase after *Rita*.

[FIGURE 2 ABOUT HERE]

Figure 2 displays average prison sentence in months for blacks and whites over the sample timeframe. There is a decline in black sentences after *Koon*, and an increase for both groups after the PROTECT Act. There is a dip prior to *Booker* which may reflect

²⁵ Although we consider Hispanics in the regression results, we only include blacks and whites in the graphs for clarity of presentation. Offenders reported as “Hispanic” were coded as Hispanic regardless of what other race was reported. Rates for Hispanics usually fell somewhere between those of whites and blacks.

the effect of *Blakely v. Washington*,²⁶ in which the Supreme Court struck down a Washington state’s sentencing regime. This may represent the deferral of harder cases until after the Supreme Court clarified the validity of the Guidelines, or courts and prosecutors anticipating the Court’s decision in *Booker*. This pre-trend might confound the *Booker* event analysis, and so we urge caution in interpreting our *Booker* results. There is a clear decrease in the prison sentences of white offenders after *Rita*, while the change in black sentences is far less pronounced.

[FIGURE 3 ABOUT HERE]

Figure 3 traces the percentage of sentences that are at the statutory minimum for blacks and whites. This percentage decreases after the introduction of the safety-valve departure in 1994, but the decrease is proportionately larger for whites. Likewise, the percentage for both groups increases significantly after *Booker*—again, with some evidence of an anticipation effect—and especially for blacks after *Rita*.

[TABLE 1 ABOUT HERE]

B. Baseline regressions

Table 1 reports the results of the baseline regression described above for the entire sample of serious crimes. Each regression includes a full set of district-year-offense interactions. This controls for district and offense-type trends and for changes in the composition of the district bench.

The point estimates suggest that under de novo review, blacks receive roughly 0.7 percentage points fewer departures than whites, 21 percent longer prison sentences, and are 0.7 percentage points more likely to be sentenced at the mandatory minimum. Black offenders also have 0.25 higher offense levels than whites. Likewise, Hispanic offenders receive 1.3 percentage points fewer departures than whites, 16 percent longer prison sentences, and are 2.8 percentage points more likely to be sentenced at the mandatory minimum. All of these differences were significant at the 1% level. In short, there are pronounced racial disparities even after controlling for base offense levels, criminal history categories, and offender demographics as well as dummies for each district-year-offense category.

The primary coefficients of interest are those on the interactions *Deferential* × *Black* and *Deferential* × *Hispanic*. For blacks and Hispanics, the disparities in departure rates are reduced to nearly zero in periods of deferential review. Moreover, the disparity in sentence length is 3.5 percentage points lower in discretionary review

²⁶ 542 U.S. 296 (2004).

periods for both blacks and Hispanics. The estimated degree of excess prison sentence for blacks falls from 21 percent under de novo review to 17.5 percent under deferential review. For Hispanics, this excess prison sentence is reduced from 16 percent to 12.5 percent. Offense levels for blacks are barely changed under deferential review, although Hispanics experience a small but statistically significant reduction in offense levels. The rate at which blacks are sentenced at the statutory minimum increases by 1.1 percentage points in periods of deferential review. The opposite appears to hold for Hispanics.

Table 1 reveals a consistent story for black offenders under deferential review standards. The disparity in departure rates and prison sentences relative to whites narrows in periods of deferential review, but black offenders are sentenced more often at the statutory minimum even as their offense levels do not change. In other words, when judges are freer to depart, they do so more proportionally more often for blacks than whites, resulting in lower prison sentences. However, judges appear to be constrained more frequently by mandatory minimums when sentencing black defendants. The story is not so consistent for Hispanics. While Hispanic-white prison disparities are lower in periods of greater judicial discretion, so are disparities in offense levels and percent at the binding minimums. Decreases in offense levels and binding minimums might indicate a decrease in severity of crimes for Hispanic offenders over the sample time frame, which could also explain the decrease in the prison sentences. We cannot draw a firm conclusion regarding the impact of deferential review on Hispanic offenders.

C. Three-Year Time Windows

We now run our baseline regression for three-year windows around the four doctrinal changes that affected sentencing discretion: (1) the *Koon* decision in June 1996, which changed the standard of review for departures from de novo to deferential; (2) the subsequent reversal of *Koon* by the PROTECT Act in March 2003, restoring de novo review; (3) the *Booker* decision in January 2005, rendering the Guidelines advisory; and (4) *Rita* in July 2007 and *Gall* and *Kimbrough* in December 2007, which shifted from the uncertain advisory Guidelines regime immediately following *Booker* to a regime in which the Supreme Court affirmed district judges' authority to sentence offenders outside the Guidelines range. We exclude the transitional six-month window between *Rita* and *Gall/Kimbrough*.

We supplement the regressions with non-parametric analysis of the impact of the doctrinal changes on the distribution of sentences. Figures 4 through 7 are kernel density plots comparing the change in the distribution of prison sentences for blacks and whites during each of the three-year windows. In each graph, the distribution of prison sentences for the 18 months prior to the doctrinal change is overlaid with the distribution of prison sentences 18 months after the change. Tables 2–5 then report regression results for the same three-year windows using the baseline regression

specification. We also present results from logistic regressions on departures, which we report in terms of odds ratios. The results from the linear probability models can be interpreted as changes in absolute disparity, while the logistic regression results can be interpreted as changes in relative disparity.

[FIGURES 4A/4B ABOUT HERE]
[TABLE 2 ABOUT HERE]

1. *Koon*.

Figure 4A and shows that the distribution of black sentences after *Koon* shifted almost uniformly to the left, with more very low sentences meted out to blacks. By contrast, Figure 4B shows that the white sentencing distribution hardly changed, with perhaps a slight increase detectable around 24–36 months. Table 2 is consistent with this result. Black departures increased by 1.1 percentage points after *Koon*, though significant at only the 10% level. However, the results from the logistic regressions suggest a stronger effect: the departure disparity narrowed by roughly 20 percent, a result that is significant at the 1% level. Prison sentences for blacks fell by 5.6 percent relative to whites. There was no significant impact on average white prison sentences, as measured by the coefficient on *Koon*. The results for Hispanics were similar, but again a drop in offense levels might indicate a trend toward less serious crimes that might independently justify the reduced prison sentences.

[FIGURES 5A/5B ABOUT HERE]
[TABLE 3 ABOUT HERE]

2. PROTECT Act.

Figures 5A and 5B suggest that the PROTECT Act did not cause an increase in disparity between blacks and whites, with both black and white sentencing distributions simply shifting to the right after the adoption of the PROTECT Act. Table 3 is consistent with Figures 5A and 5B. While the absolute departure disparity, reported in the linear probability model, was lower during the PROTECT period, there were also far fewer departures for both groups after the PROTECT Act. There was no significant change in the relative racial disparity in departure rates, reported in the logistic regression. The disparity in prison sentences was also unchanged.

If the PROTECT Act merely reverted to the pre-*Koon* de novo standard, one would expect a reversion back to pre-*Koon* levels of racial disparities. Instead, we find no measureable impact of the PROTECT Act. However, the PROTECT Act mandated the strictest standard of review ever applied to departure decisions, and the Commission placed a moratorium on new grounds for departures. Therefore, the

PROTECT Act era was likely the period during which federal sentencing was most constrained, and it is possible that the move toward more determinate sentencing may not have increased racial disparities. At the very least, the results from the PROTECT Act event window are consistent with the notion that judicial bias was not a significant source of racial disparity under the Guidelines. A significant reform limited judicial discretion, increased sentences, and reduced departures, but there was no statistically significant change in racial disparity.

[FIGURES 6A/6B ABOUT HERE]
[TABLE 4 ABOUT HERE]

3. *Booker*.

Figures 6A and 6B demonstrate that black and white sentencing distributions changed little after *Booker*, except that sentences became slightly more heaped around the mandatory minimums for both races. As we cautioned earlier, it is possible that there were significant anticipation effects prior to *Booker*, with more serious cases being delayed. Table 4, which controls for offense and offender characteristics, suggests a more substantial change. After *Booker*, departures increased, prison sentences decreased, and the percentage of sentences at the binding minimum increased. However, despite these significant shifts, there were no changes in the relative disparities among blacks, whites, and Hispanics in any of our outcome variables. Interestingly, there is a 0.18 increase in offense level calculations.²⁷ We caution again that the impact of *Booker* is potentially confounded by the anticipation of *Booker* following *Blakely*.

[FIGURES 7A/7B ABOUT HERE]
[TABLE 5 ABOUT HERE]

4. *Rita, Gall, and Kimbrough*.

Figures 7A and 7B, which depict the impact of *RGK* on the distribution of sentences, reveal a stark contrast in the effects on black and white defendants. The distribution of black prison sentences becomes heaped around the standard statutory minimums of 5, 10, 15, and 20 years, while white prison sentences exhibit a more or less uniform shift to the left. Table 5 is consistent with this finding. Departures increase after *RGK*, but proportionately less for blacks. This does not necessarily constitute evidence of discrimination because departures are sometimes foreclosed by mandatory

²⁷ One interesting speculation would be that judges did not need to use offense level adjustments to reach their preferred sentences, but could rely more heavily on departures, as evidenced by the 14 percentage point increase in the departure rate after *Booker*. It is also possible, however, that prosecutors argued for more enhancements or charged offenders with more severe crimes.

minimums, which may apply more often to black offenders. Prison sentences decrease for whites, but are basically unchanged for blacks; the coefficients on *RGK* (-9.5%) and *RGK* × *Black* (12.5%) are roughly offsetting. Moreover, after *RGK*, blacks are sentenced at the binding minimum 3 percentage points more often than whites, and 2 percentage points more often for minimums of 10 years or more. In sum, the sentencing distribution for black offenders shifts somewhat to the left after *RGK*, but part of the shift that would otherwise have occurred is impeded by the mandatory minimums. The disparity between white and black offenders in prison sentences increases, but the weight of the evidence suggests that this is largely a consequence of the mandatory minimums. Sentences for crack cocaine trafficking were reduced through lowered base offense levels in 2007, however, we excluded crack cases in unreported regressions, and our results were similar.

Interestingly, Hispanic prison sentences declined after *RGK* by roughly the same amount as white prison sentences. There is little evidence that *RGK* increased disparities between white and Hispanic offenders, even though Hispanics and blacks had roughly the same level of unexplained prison sentence disparity prior to *RGK*. The key again appears to be mandatory minimums. After *RGK*, there was little change in the proportion of Hispanic offenders sentenced at the statutory minimum.

There is some evidence of changes in offense level calculations around *RGK* and *Booker*, increasing after *Booker* but decreasing by roughly the same amount after *RGK*. This may reflect changes in both judicial and prosecutorial behavior, and reinforces our concern that base offense level is endogenous to the sentencing regime and therefore not an appropriate control. Regressions taking base offense levels as the dependent variable also suggest that base offense levels may have changed. We take account of this possibility in the next analysis.

[TABLE 6 ABOUT HERE]

D. Mandatory Minimums after Rita, Gall, and Kimbrough

We have thus far presented only suggestive evidence that the increasing importance of mandatory minimums after *RGK* drove much of the increase in racial disparities observed. Table 6 tests this hypothesis directly by dividing the sample into two groups: cases for which mandatory minimums are less likely to be binding and those for which minimums are more likely to apply. The former category consists of cases involving offenders in the lowest criminal history category who did not use firearms. The latter category consists of offenders with substantial criminal histories and

cases involving firearms offenses.²⁸ For the sample as a whole, 9 percent of cases in the “likely binding” categories were sentenced at the mandatory minimum versus 6 percent of the remaining cases. In the period after *RGK*, this level increased to 13 percent of cases in the “likely binding” categories but only 7 percent for the remaining cases.

The columns 1–4 report the impact of *RGK* on cases for which mandatory minimums are less relevant, and columns 5–8 report cases for which mandatory minimums are more likely to be binding. We report results both including and excluding base offense level as a control variable. Base offense levels may be an important control because they reflect offense severity. However, charging behavior appears to have changed following *Booker* and *RGK*, which means that base offense levels may themselves have been affected by these reforms. We accommodate both possibilities.

In the subset of cases in which minimums are less likely to be relevant, the proportion of offenders sentenced at the minimum increases slightly after *RGK*, but the effect is the same for whites and blacks. The effect on prison sentences depends on whether we control for base offense levels. In the regressions that do not control for base offense levels (Column 3), there is a decrease in prison sentences post-*RGK* and the decrease for whites and blacks is equal. This is a marked contrast to our previous results that there is a decrease for whites but not for blacks. When we condition on base offense levels (Column 4), white prison sentences decrease by 13 percent, but the coefficient on the interaction term between black and *RGK* is a positive 8.4 percent, which means that sentences were not reduced to the same extent for blacks. The point estimates imply that the decrease for black sentences was about one-third that of whites, and the change in black sentences overall was not statistically significant (–4.6 percent with a *p*-value of 0.18). Turning to crimes for which mandatory minimums should be relatively more important, blacks are 3.4 percentage points more likely to be sentenced at the binding minimum after *RGK*. Moreover, the decrease in prison sentences for whites is not shared by blacks. In Column 6, which does not control for base offense levels, white sentences declined by 9.4 percent. The coefficient of 7.5 percent on the interaction between black and *RGK*, however, implies that there was almost no change in black sentences after *RGK*. The same result obtains when we condition on base offense levels, with white prison sentences declining by 11 percent and the black interaction term of 10 percent indicating that black sentences do not change in a statistically or qualitatively meaningful sense.

²⁸ This division is defensible because the categories are likely to be exogenous. Criminal history is largely pre-determined by prior convictions, and hence should be unaffected by doctrinal changes and not subject to manipulation. Within broad categories of offenses, such as firearms and drug trafficking, there should not be much substitution between categories.

One additional concern with a before-after comparison is that there may be pre-existing race-specific trends that bias estimation. To check for this, we divide the sample into blacks and whites and separately estimate the pre- and post-RGK trends for each racial group. We take the six months between *Rita* and *Gall/Kimbrough* as the comparison (or excluded) period, and report the coefficients on the time dummies for the six other six-month periods before and after *Rita* in figures 8 through 10. These graphs present a clear picture. Departures substantially increase for both whites and blacks after *RGK*, but prison sentences only decline for whites. Black prison sentences remain relatively flat, while the percentage of sentences at the binding minimum increase during the six-month window after *Rita* by about 2 percentage points and by another two percentage points after *Gall* and *Kimbrough*. The departure results suggest judges prefer more lenient prison sentences for black offenders, but the binding minimum graph shows why this may not have translated into lower prison sentences overall.

V. Conclusion

The United States Sentencing Guidelines were motivated in part by a belief that restricting judicial discretion would mitigate unwarranted racial disparities in criminal sentencing. However, the focus of the Guidelines on judicial discretion ignores the effect of prosecutors and other institutional factors on sentencing. By examining changes in the standard of review applied to judges' departures from the Guidelines, we conclude that racial disparities can actually be reduced by judicial discretion, at least in the context of Guidelines sentencing.

On average, racial disparities are smaller during periods of deferential review under the *Koon* and *Booker* periods. More detailed examinations of three-year event windows find that racial disparities were reduced after *Koon*, and perhaps little changed after the PROTECT Act and *Booker*. Another striking finding is that racial disparities were little changed by the PROTECT Act, even though the Act adopted strict review standards, encouraged prosecutorial appeals of departures, and limited judicial discretion to make new grounds for departures. Comparing the PROTECT Act period to either the period before it was enacted or the period after *Booker*, racial disparities remained roughly constant even as overall prison sentences and departure rates were affected.

Simplistic comparisons suggests that racial disparity increased after recent Supreme Court decisions in *Rita*, *Gall*, and *Kimbrough* clarified *Booker* and encouraged further departures from the Guidelines. Most of the post-RGK increase in disparity, however, is due to the increased relevance of statutory minimums under a system of advisory Guidelines. White and Hispanic prison sentences declined after *Rita*. Black prison sentences did not, except in those cases in which mandatory minimums were not

likely to bind. The most straightforward explanation for these results is that mandatory minimum sentences were more constraining for black offenders. Direct evidence of this is provided by the proportion of sentences at the mandatory minimum, and the fact that the effect of *Rita* for blacks and whites is more similar when mandatory minimums are less binding.

We cannot confidently say what racial disparities would be in the absence of mandatory minimums. Nonetheless we have some evidence that when mandatory minimums are less relevant, sentences fell by the same proportion for whites and blacks. In short, our findings suggest that judicial discretion does not contribute to, and may in fact mitigate, racial disparities in Guidelines sentencing. Policy makers interested in redressing racial disparity today should pay much closer attention to the effects of mandatory minimums and their effect on prosecutorial and judicial discretion.

Bibliography

Albonetti, Celesta A. (1997) "Sentencing under the Federal Sentencing Guidelines: Effects of Defendant Characteristics, Guilty Pleas, and Departures on Sentence Outcomes for Drug Offenses," 31 *L. & Soc. Rev.* 789.

Alexander, Michelle (2010) *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: New Press.

Anderson, John M., Jeffrey R. Kling, & Kate Stith (1999) "Measuring Interjudge Sentencing Disparity: Before and After the Federal Sentencing Guidelines," 42 *J. of L. & Economics* 271.

Bibas, Stephanos (2001) "Judicial Fact-Finding and Sentence Enhancements in a World of Guilty Pleas," 110 *Yale L. J.* 1097.

Baron-Evans, Amy, & Kate Stith (2012) "Booker Rules," 160 *Univ. of Pennsylvania L. Rev.* 1631.

Bowman, Frank O., & Michael Heise (2001) "Quiet Rebellion? Explaining Nearly a Decade of Declining Federal Drug Sentences," 86 *Iowa L. Rev.* 1043.

Bushway, Shawn D., & Anne Morrison Piehl (2001) "Judging Judicial Discretion: Legal Factors and Racial Discrimination in Sentencing," 35(4) *L. & Soc. Rev.* 733.

Everett, Ronald S., & Roger A. Wojtkiewicz (2002) "Difference, Disparity, and Race/Ethnic Bias in Federal Sentencing," 18 *J. of Quantitative Criminology* 189.

Etienne, Margareth (2006) "Pain and Race: A New Understanding of Race Based Sentencing Disparities," 3 *Univ. of St. Thomas L. J.* 496.

Fischman, Joshua B., & Max M. Schanzenbach (2011) "Do Standards of Review Matter? The Case of Federal Criminal Sentencing," 40 *J. of Legal Studies* 405.

Freeborn, Beth A., & Monica E. Hartmann (2010) "Judicial Discretion and Sentencing Behavior: Did the Feeney Amendment Rein in District Judges?" 7(2) *J. of Empirical Legal Studies* 355.

Hofer, Paul J. (2012) "Review of the U. S. Sentencing Commission's Report to Congress: Mandatory Minimum Penalties in the Federal Criminal Justice System," 24(3) *Federal Sentencing Reporter* 193.

Hofer, Paul J., Kevin R. Blackwell, & R. Barry Ruback (1999) "The Effect of the Federal Sentencing Guidelines on Inter-Judge Sentencing Disparity," 90 *J. of Criminal L. and Criminology* 239.

Johnson, Barry L. (1998) "Discretion and the Rule of Law in Federal Guidelines Sentencing: Developing Departure Jurisprudence in the Wake of *Koon v. United States*," 58 *Ohio State L. J.* 1697.

LaCasse, Chantale, & A. Abigail Payne (1999) "Federal Sentencing Guidelines and Mandatory Minimum Sentences: Do Defendants Bargain in the Shadow of the Judge?" 42 *J. of L. & Economics* 245.

Lu, Lynn D. (2007) "Prosecutorial Discretion and Racial Disparities in Federal Sentencing: Some Views of Former U.S. Attorneys," 19(3) *Federal Sentencing Reporter* 192.

Maxfield, Linda Drazga, & John H. Kramer (1998) *Substantial Assistance: An Empirical Yardstick Gauging Equity in Current Federal Policy and Practice*. Washington, DC: United States Sentencing Commission.

McDonald, Douglas C., & Kenneth E. Carlson (1994) "Why Did Racial/Ethnic Sentencing Differences in Federal District Courts Grow Larger Under the Guidelines?" 6(4) *Federal Sentencing Reporter* 223.

Meierhoefer, Barbara A. (1992) *The General Effect of Mandatory Minimum Prison Terms: A Longitudinal Study of Federal Sentences Imposed*, Washington, DC: Federal Judicial Center.

Mustard, David B. (2001) "Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts," 44(1) *J. of L. & Economics* 285.

Provine Doris Marie (2007) *Unequal Under Law: Race and the War on Drugs*. Chicago: Univ. of Chicago Press.

Roberts, Dorothy E. (2004) "The Social and Moral Costs of Mass Incarceration in African American Communities" 56 *Stanford L. Rev.* 1272.

Schanzenbach, Max (2005) "Racial and Sex Disparities in Prison Sentences: The Effect of District-Level Judicial Demographics," 34 *J. of Legal Studies* 57.

Schanzenbach, Max, & Emerson Tiller (2007) "Strategic Judging under the United States Sentencing Guidelines: Positive Political Theory and Evidence," 23 *J. of L., Economics, & Organization* 24.

Schulhofer, Stephen J., & Ilene H. Nagel (1997) "Plea Negotiations Under the Federal Sentencing Guidelines: Guideline Circumvention and its Dynamics in the Post-Mistretta Period," 91 *Northwestern Univ. L. Rev.* 901.

Scott, Ryan W. (2010) "Inter-Judge Sentencing Disparity after *Booker*: A First Look," 63 *Stanford L. Rev.* 1.

Scott, Ryan W. (2011) "Race Disparity under Advisory Guidelines: Dueling Assessments and Potential Responses," 10(4) *Criminology & Public Policy* 1129.

Sklansky, David S. (1995) "Cocaine, Race, and Equal Protection," 47 *Stanford L. Rev.* 1283.

Steffensmeier, Darrell, Jeffrey Ulmer, & John Kramer (1998) "The Interaction Of Race, Gender, And Age In Criminal Sentencing: The Punishment Cost Of Being Young, Black, And Male," 36 *Criminology* 763.

Stith, Kate, & José A. Cabranes (1998) *Fear of Judging: Sentencing Guidelines in the Federal Courts*. Chicago: Univ. of Chicago Press.

Stith, Kate, & Steve Y. Koh (1993) "The Politics of Sentencing Reform: The Legislative History of the Federal Sentencing Guidelines," 28 *Wake Forest L. Rev.* 223.

Tonry, Michael (2012) "Race, Ethnicity, and Punishment," in Joan Petersilia and Kevin R. Reitz (eds.), *The Oxford Handbook of Sentencing and Corrections*. New York: Oxford Univ. Press.

Ulmer, Jeffery T., Michael T. Light, & John H. Kramer (2011) "Racial disparity in the wake of the *Booker/Fanfan* decision: An alternative analysis to the USSC's 2010 report," 10(4) *Criminology & Public Policy* 1077.

United States Sentencing Commission (2006) *Final Report of the Impact of United States v. Booker on Federal Sentencing*. Washington, DC: U.S. Sentencing Commission.

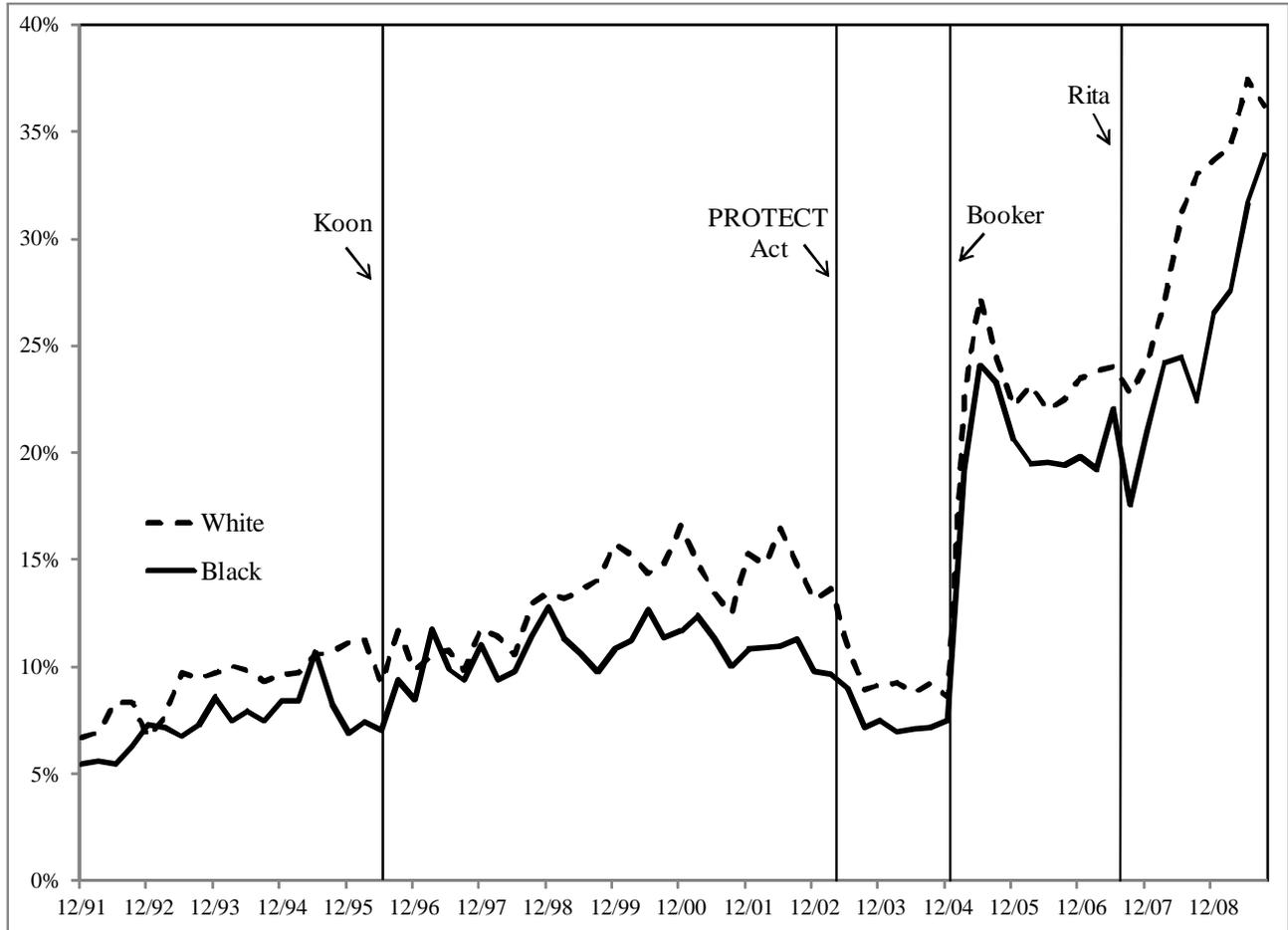
United States Sentencing Commission (2007) *Report to the Congress: Cocaine and Federal Sentencing Policy*. Washington, DC: U.S. Sentencing Commission.

United States Sentencing Commission (2010) *Demographic Differences in Federal Sentencing Practices: An Update of the Booker Report's Multivariate Regression Analysis*. Washington, DC: U.S. Sentencing Commission.

United States Sentencing Commission (2011) *Mandatory Minimum Penalties in the Federal Criminal Justice System*. Washington, DC: U.S. Sentencing Commission.

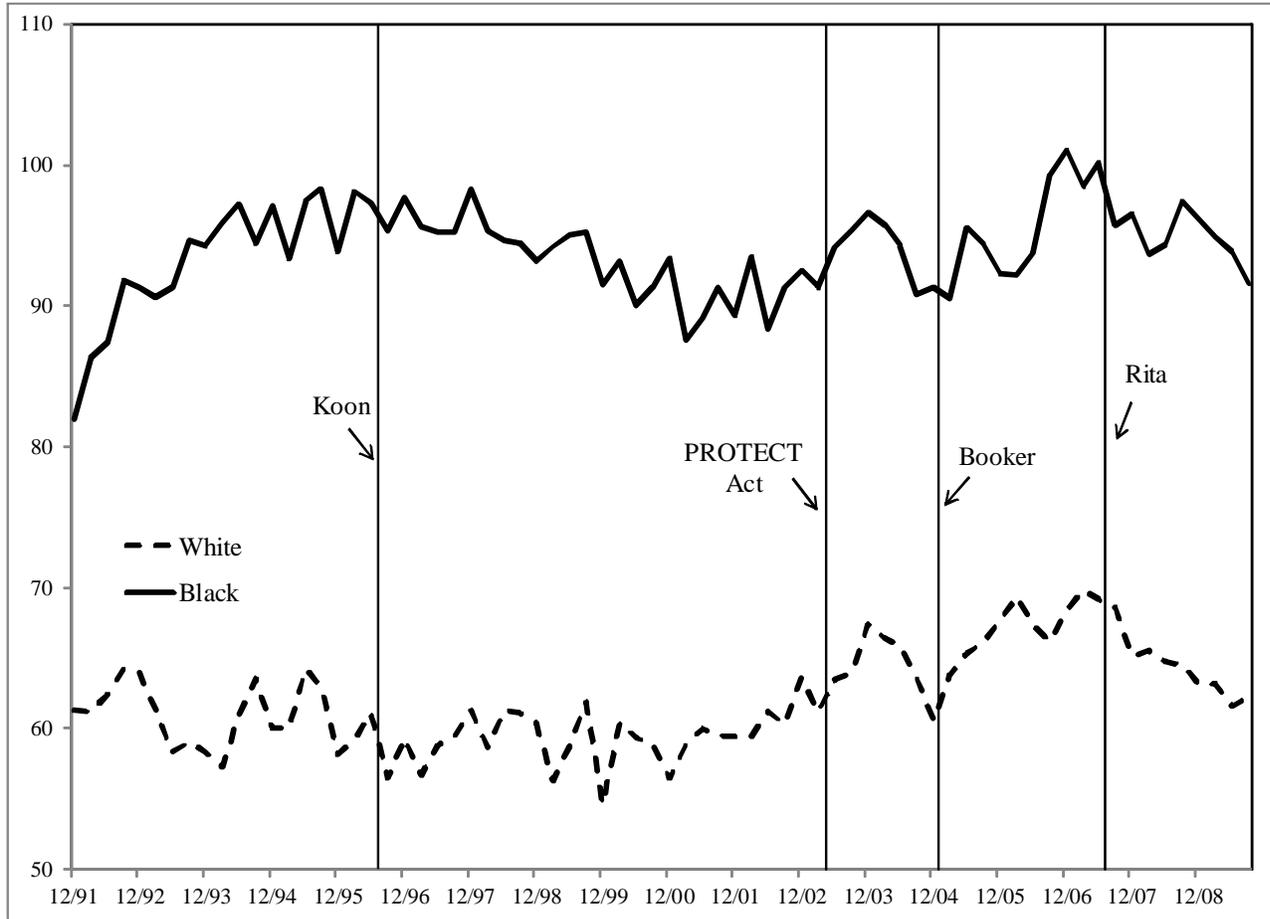
Wang, Xia, Daniel P. Mears, Cassia Spohn, & Lisa Dario (2009) "Assessing the Differential Effects of Race and Ethnicity on Sentence Outcomes Under Different Sentencing Systems," *Crime & Delinquency*, in press.

Figure 1: Judge-Induced Downward Departures, by Race



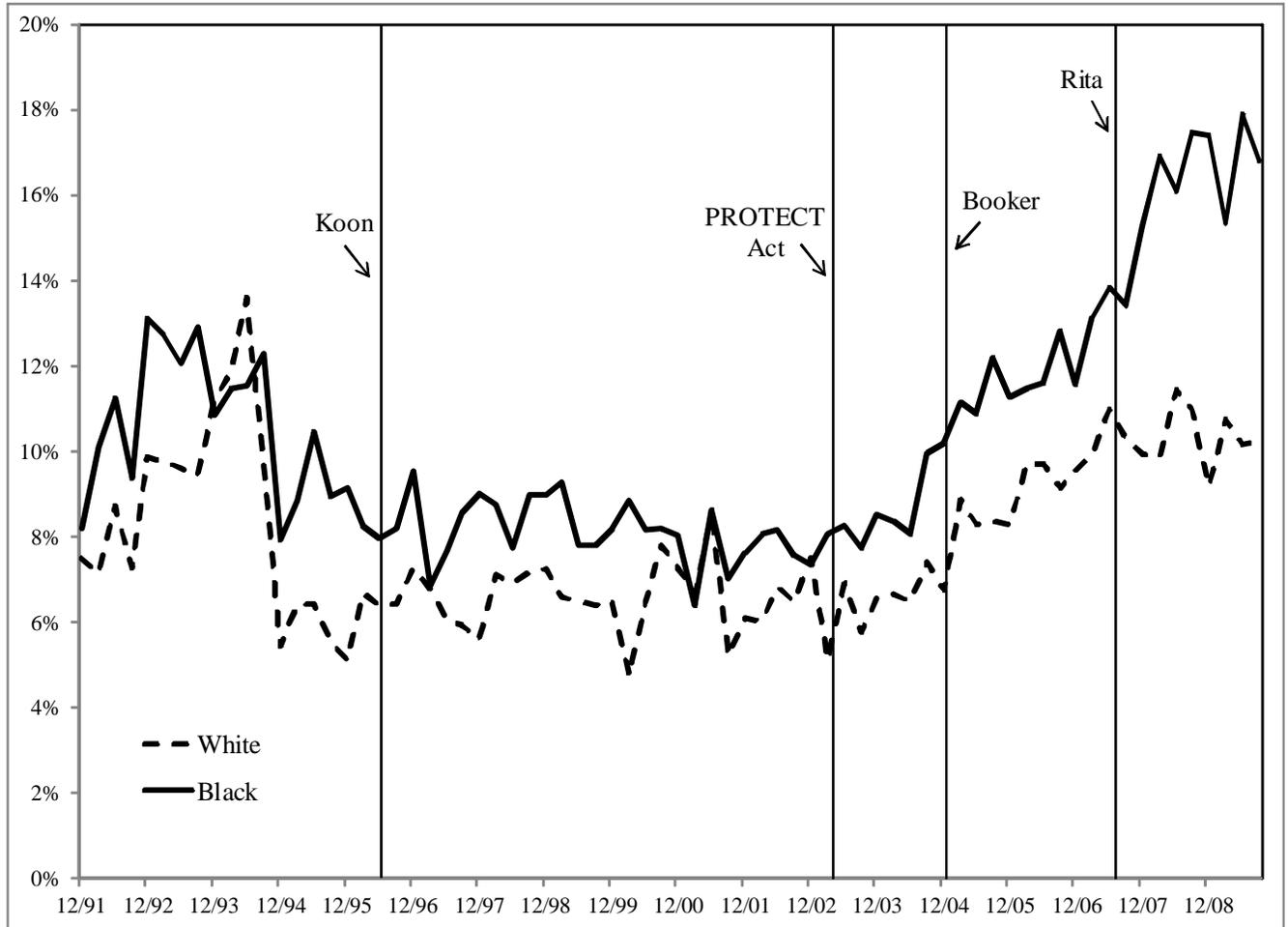
Notes: Rates of judge-induced downward departures for all serious crimes; substantial assistance departures are excluded. Rates are estimated in three-month intervals.

Figure 2: Average Prison Sentence in Months, by Race



Notes: Average prison sentence, in months, for all serious crimes. Rates are estimated in three-month intervals.

Figure 3: Offenders Sentenced at the Mandatory Minimum, by Race



Notes: Proportion of offenders sentenced at the statutory minimum, for all serious crimes. Rates are estimated in three-month intervals.

Figure 4A: Three-Year Window around *Koon*: Sentence Distribution for Black Offenders



Figure 4B: Three-Year Window around *Koon*: Sentence Distribution for White Offenders

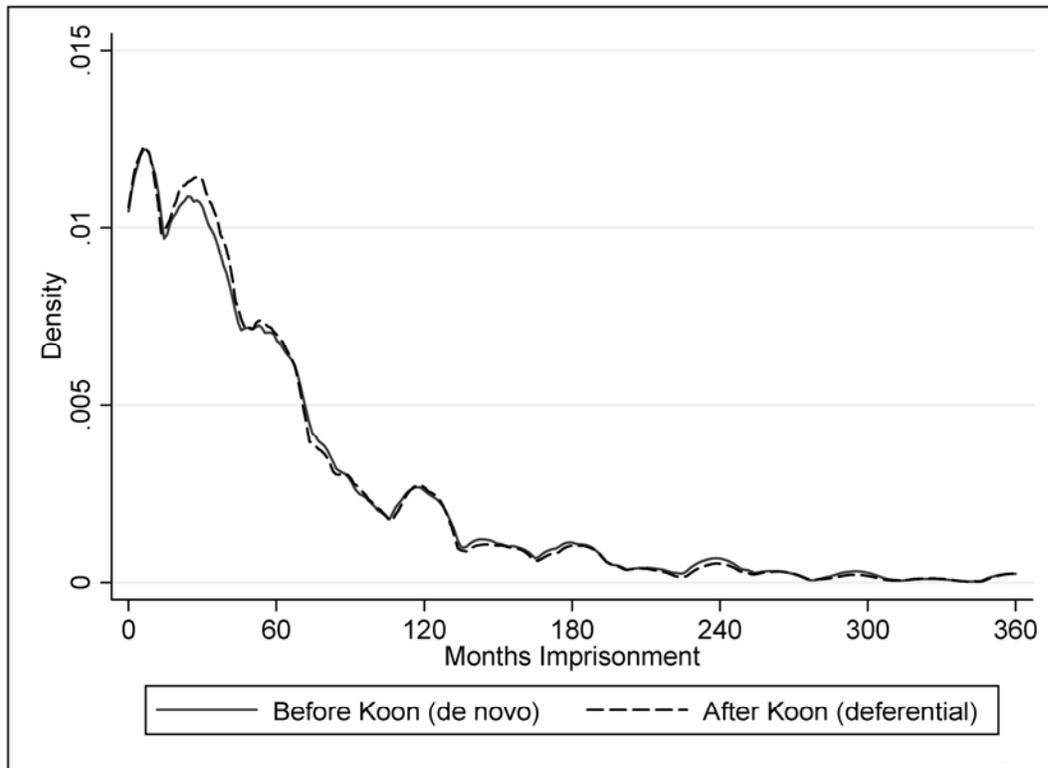


Figure 5A: Three-Year Window around PROTECT Act: Sentence Distribution for Black Offenders

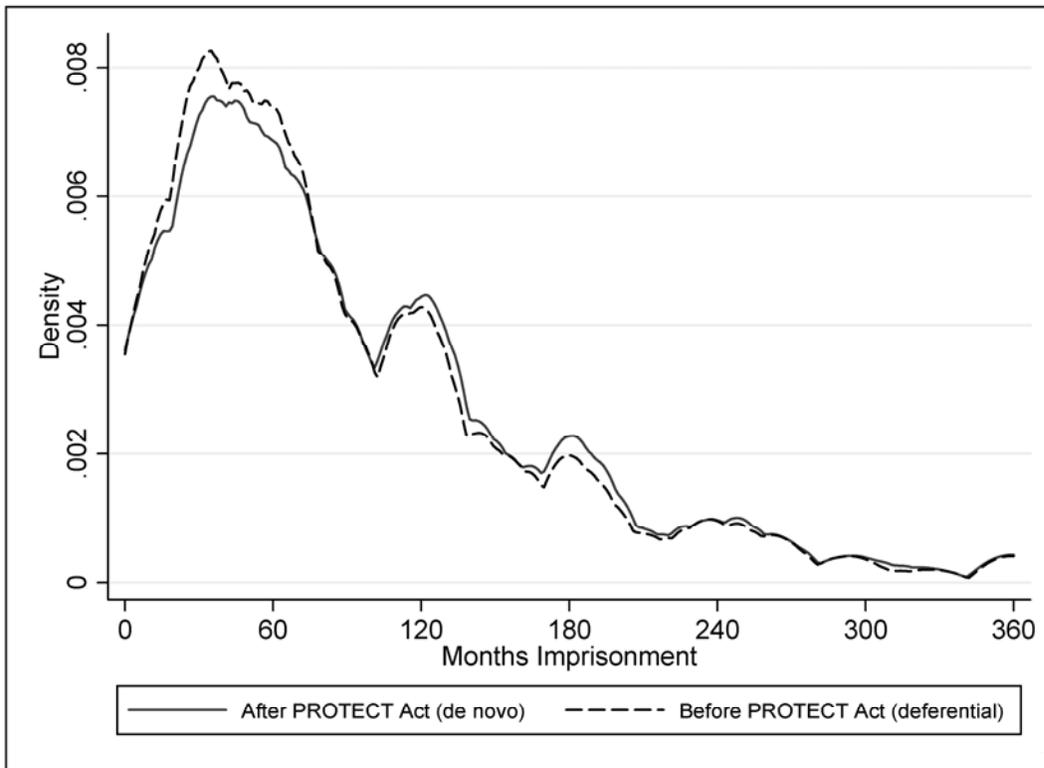


Figure 5B: Three-Year Window around PROTECT Act: Sentence Distribution for White Offenders

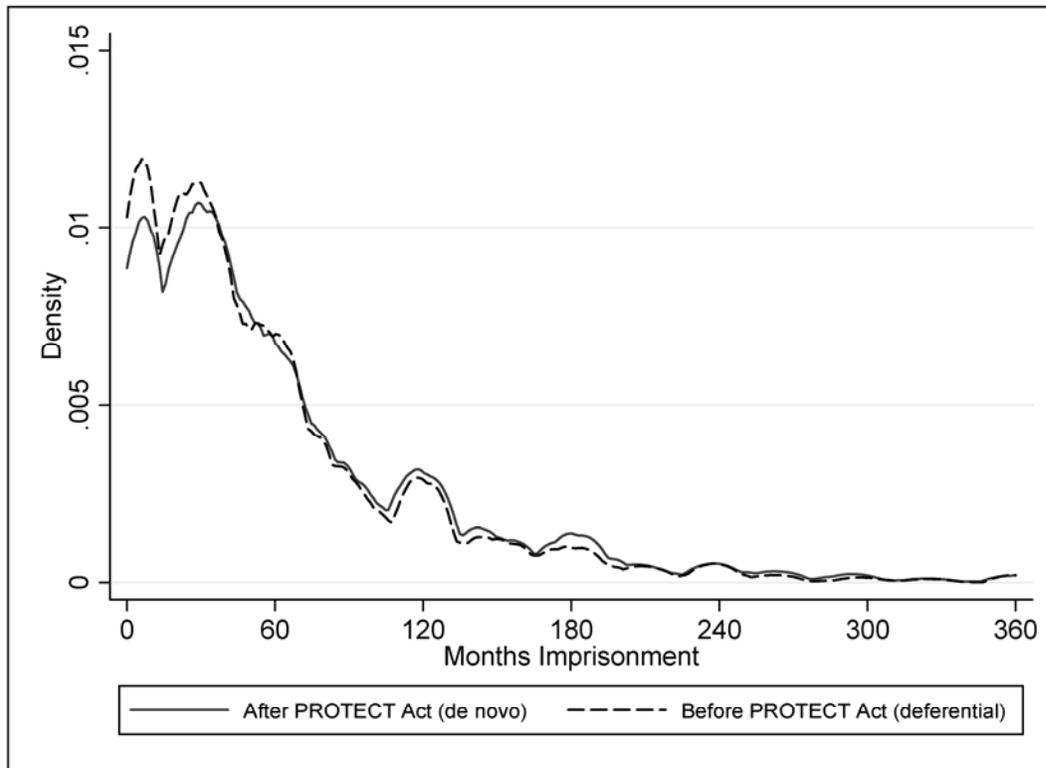


Figure 6A: Three-Year Window around *Booker*: Sentence Distribution for Black Offenders

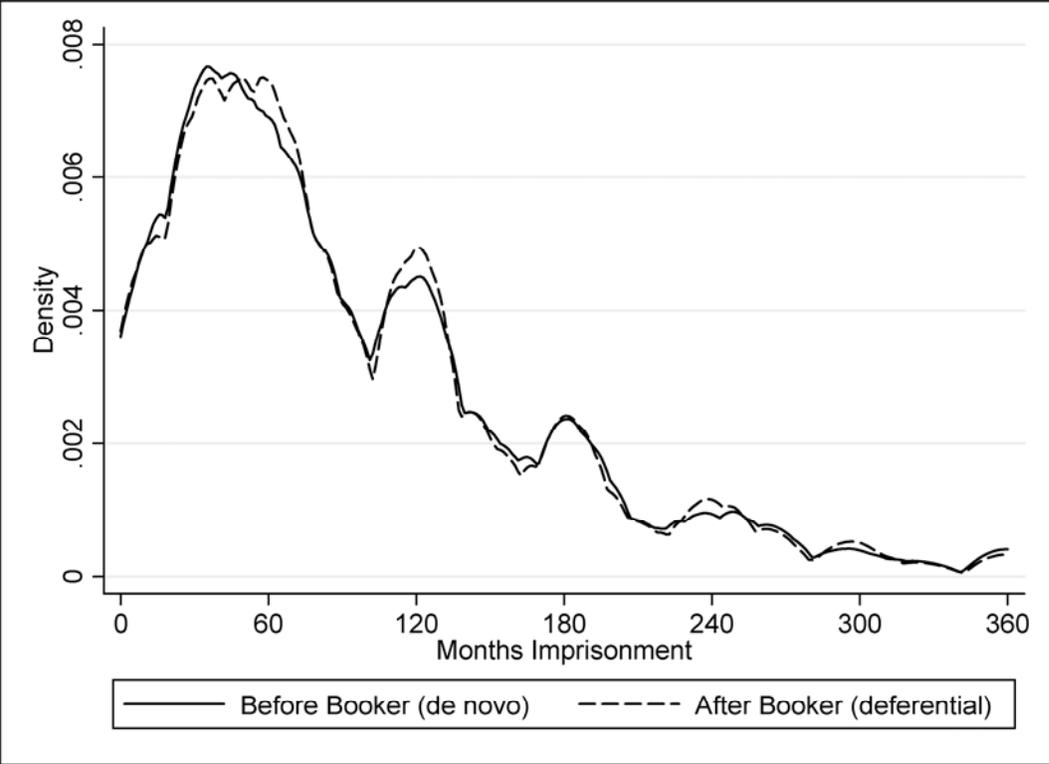


Figure 6B: Three-Year Window around *Booker*: Sentence Distribution for White Offenders

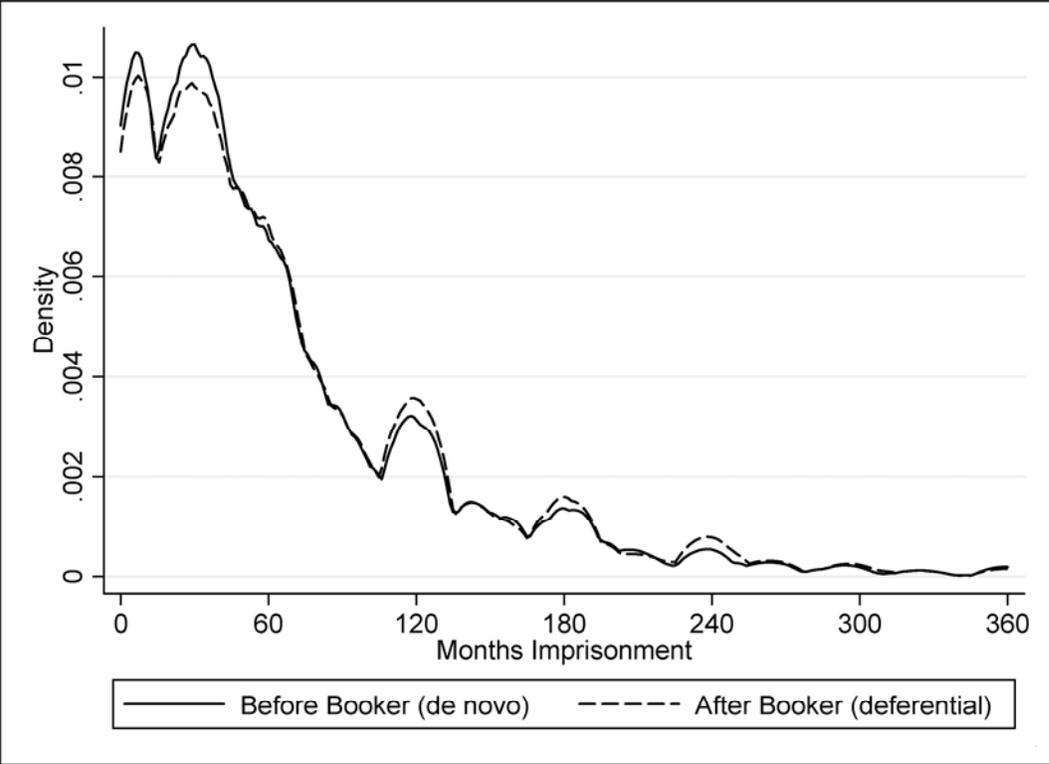
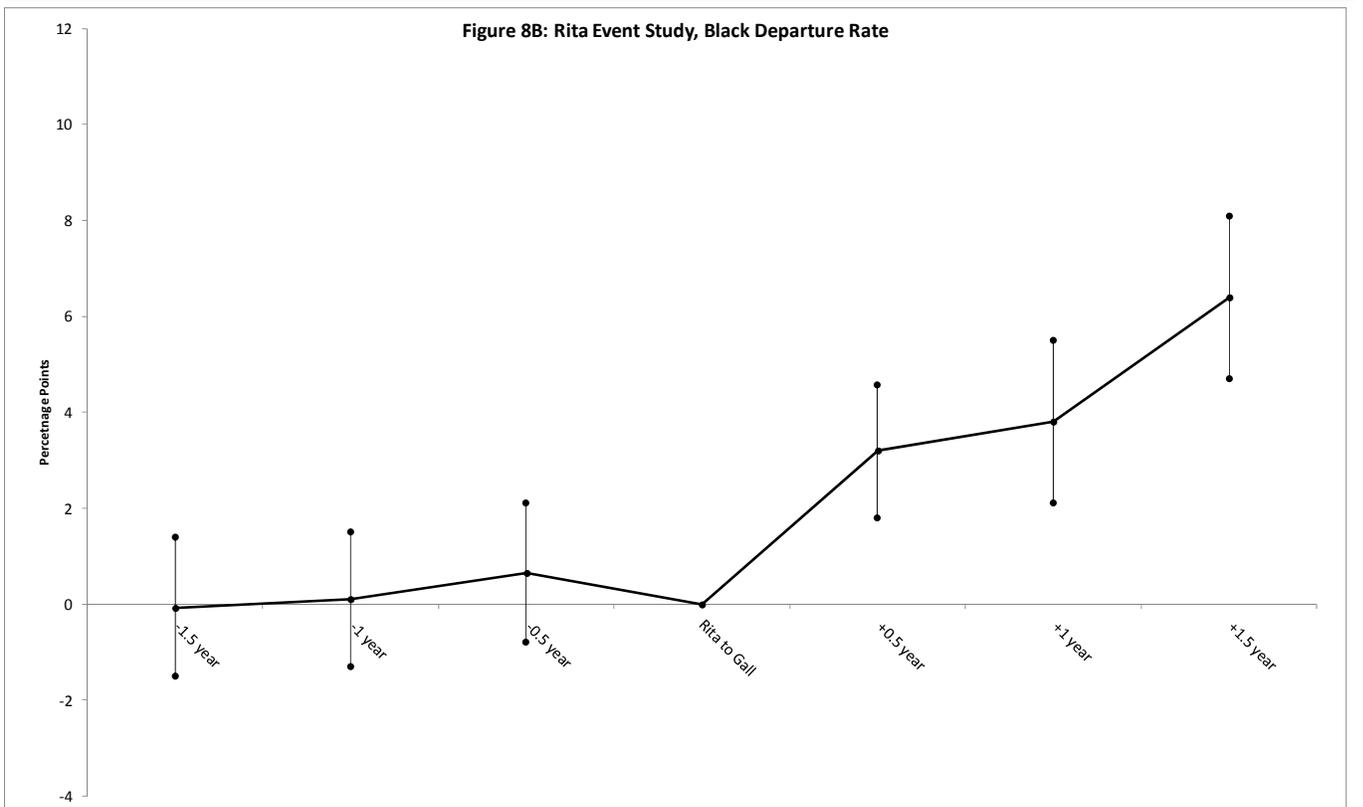
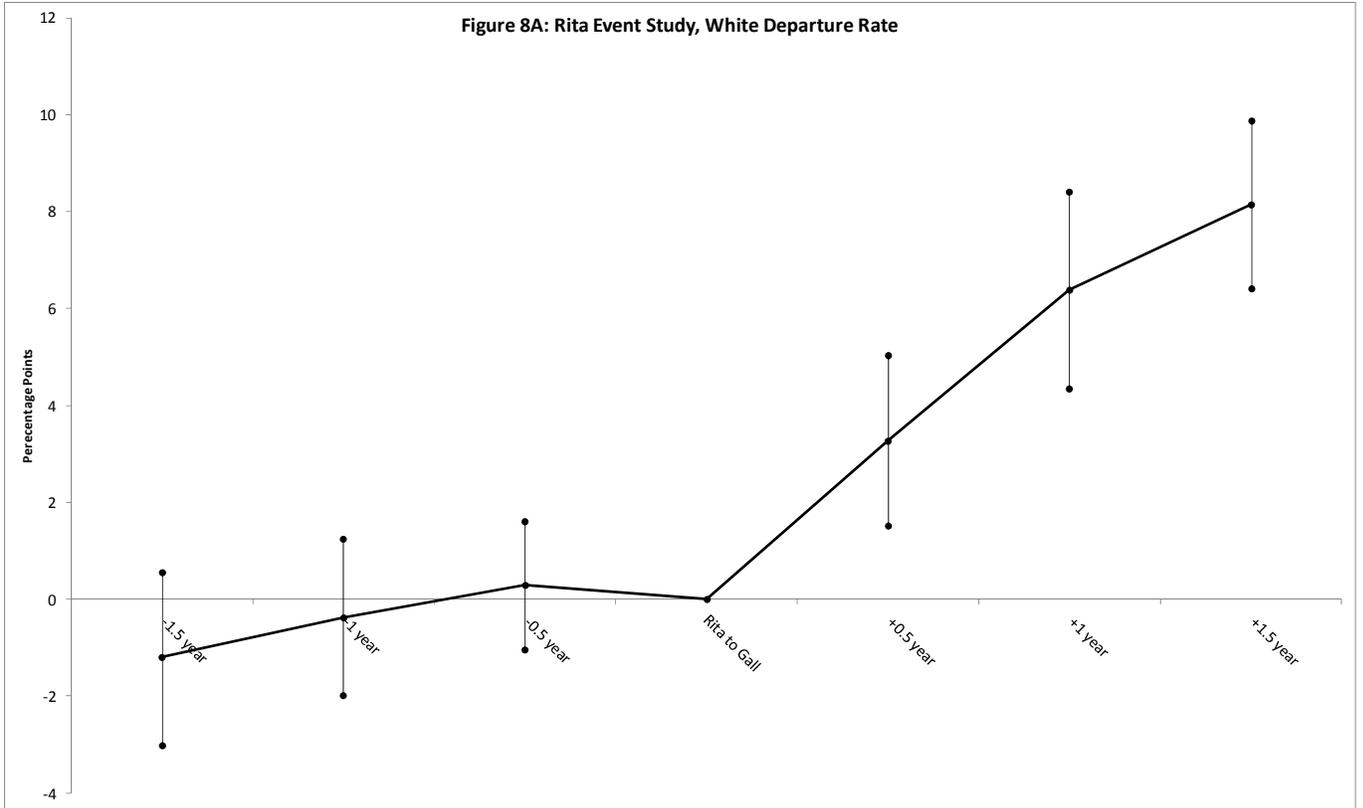


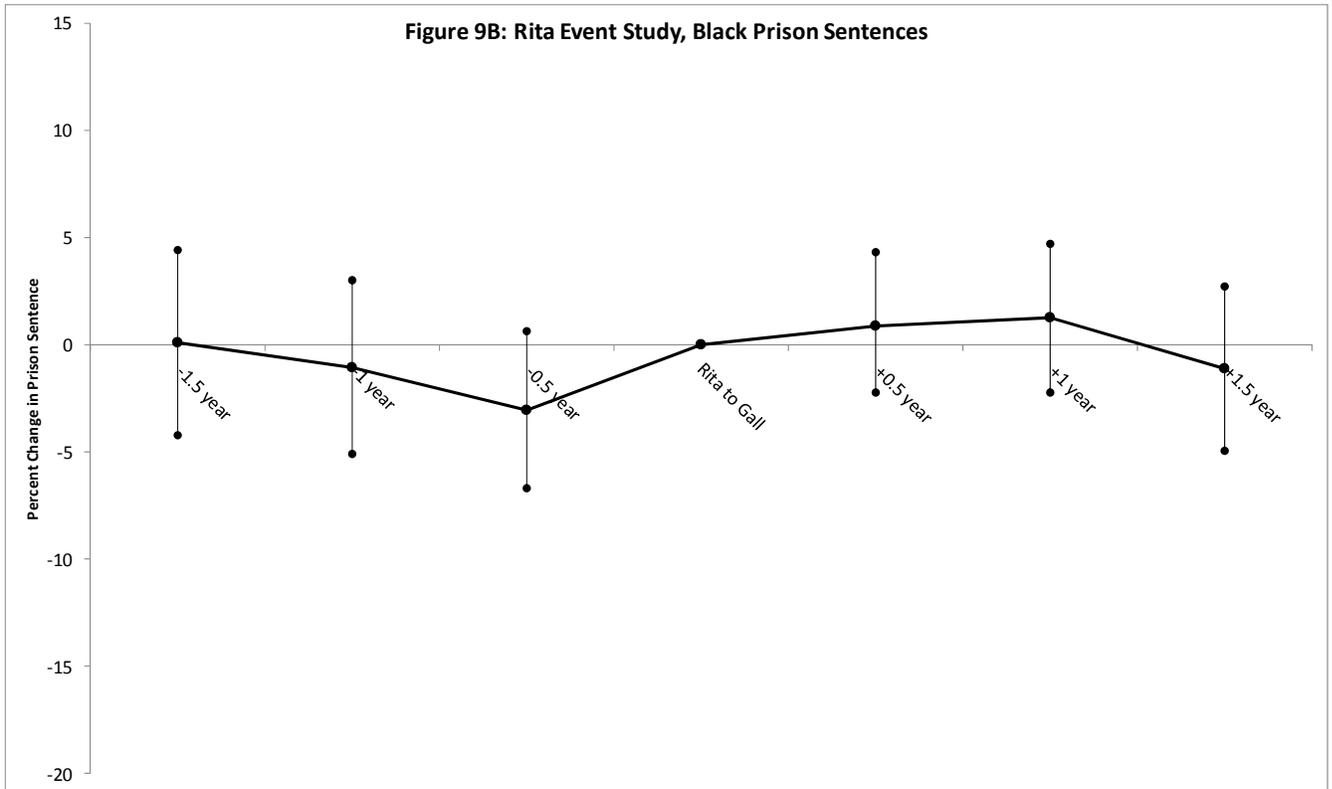
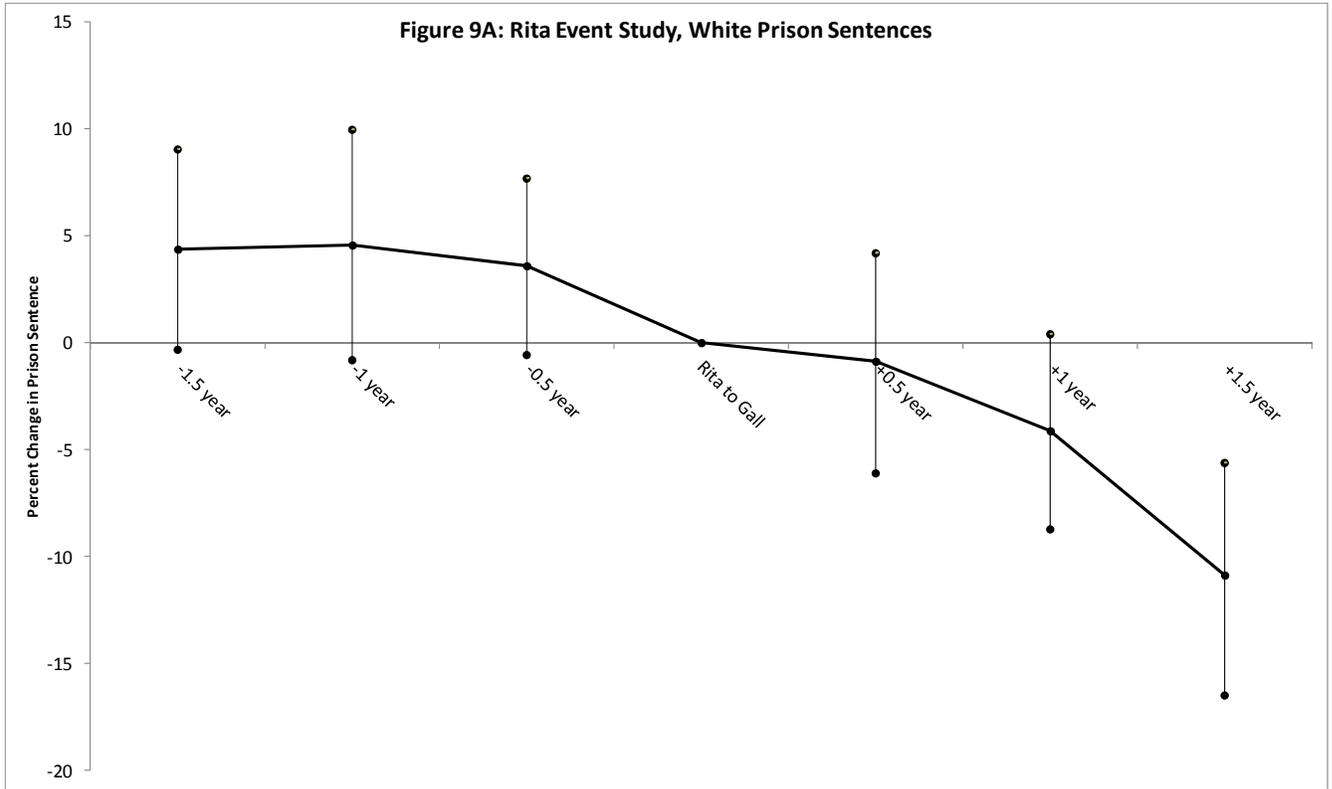
Figure 7A: Three-Year Window around *Rita*, *Gall*, and *Kimbrough*: Sentence Distribution for Black Offenders



Figure 7B: Three-Year Window around *Rita*, *Gall*, and *Kimbrough*: Sentence Distribution for White Offenders







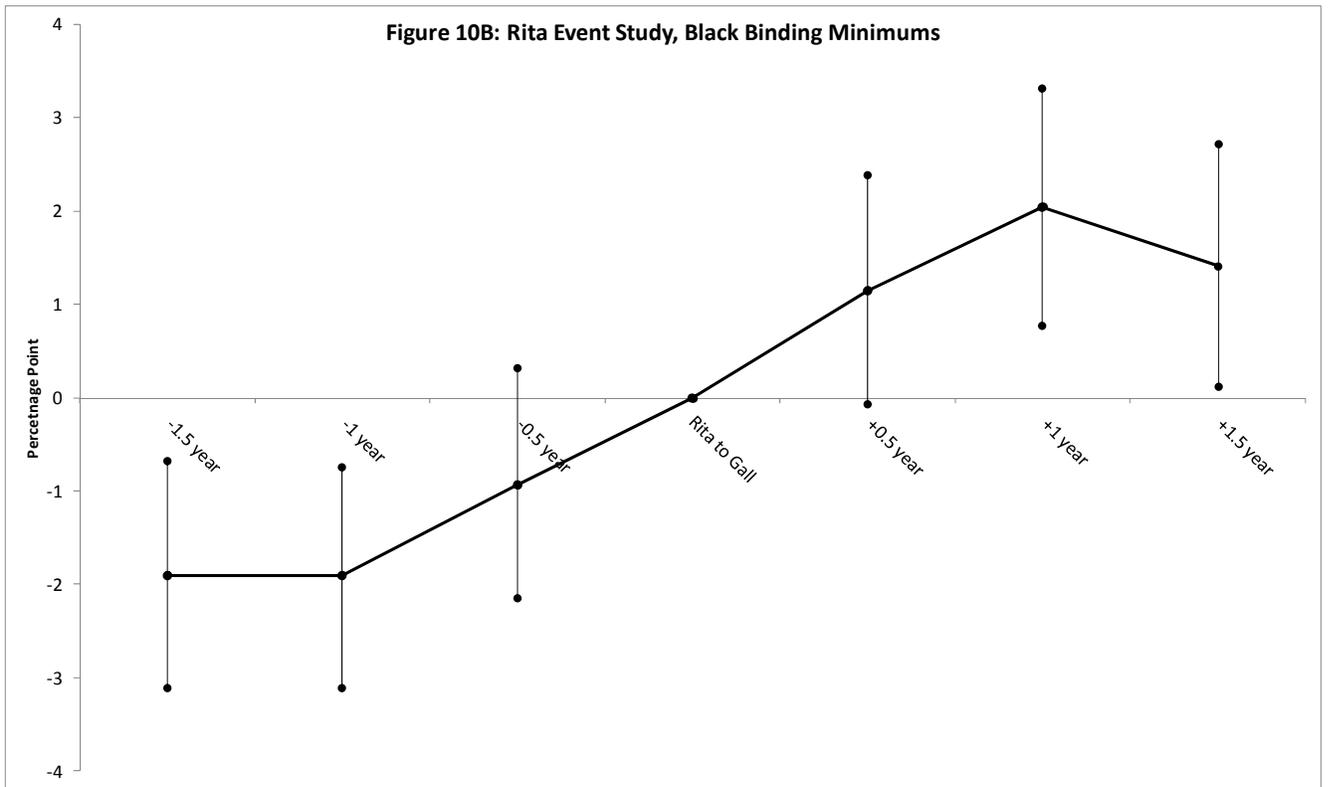
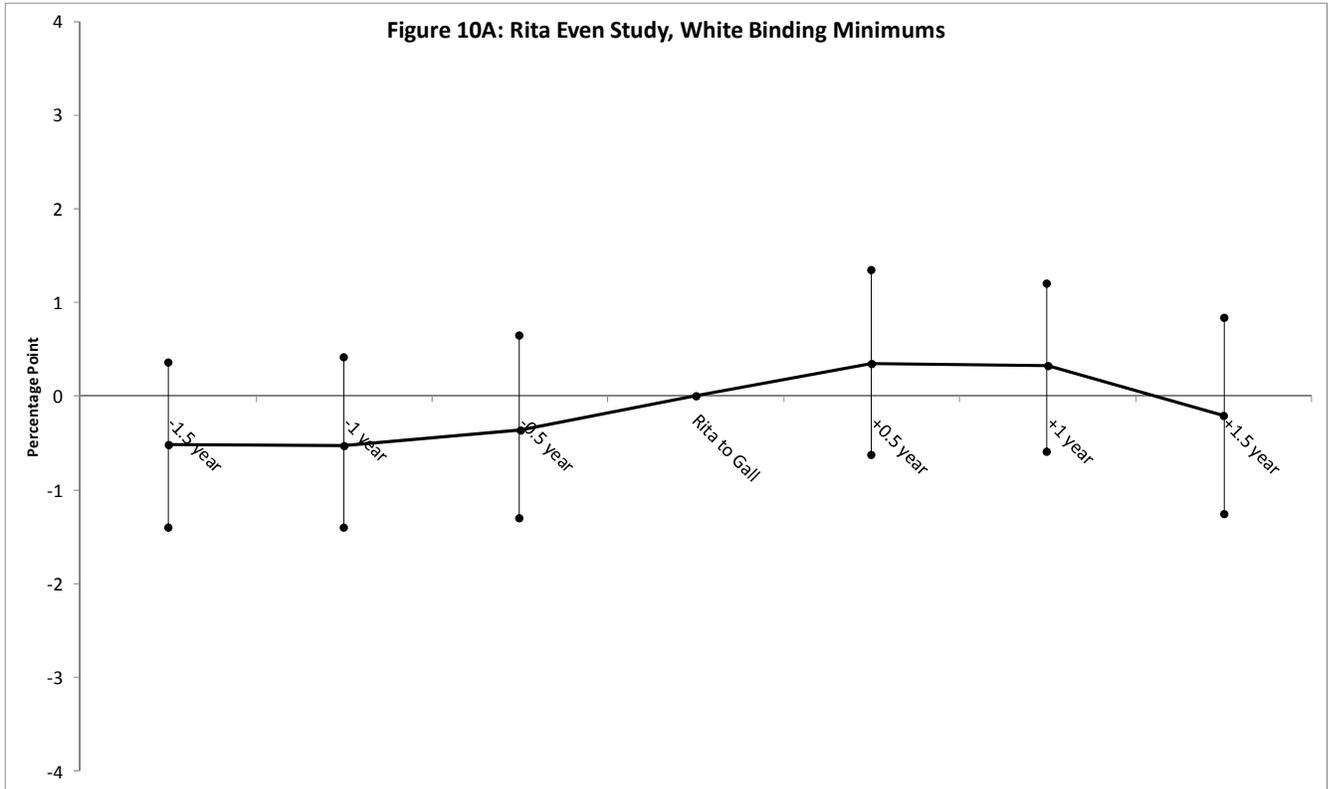


Table 1: Racial Disparities and Standards of Review

	Departure (Linear Probability)	Log Prison	Offense Level	Binding Minimum
Black Defendant	-0.007** (0.002)	0.21** (0.006)	0.25** (0.020)	0.007** (0.002)
Hispanic Defendant	-0.013** (0.002)	0.16** (0.007)	-0.026 (0.023)	0.003** (0.0002)
Deferential × Black	0.006** (0.002)	-0.035** (0.008)	-0.031 (0.023)	0.011** (0.003)
Deferential × Hispanic	0.018** (0.003)	-0.035** (0.009)	-0.15* (0.025)	-0.012** (0.004)
R²	0.13	0.90	0.89	0.068

**significant<.01; *significant at <.05; +significant<.10. All regressions include year-deferential-district-offense category interactions as well as controls for criminal history category, log of base offense level, the interaction between criminal history and log base offense level, age, age-squared, number of dependents (dummies for one, two, and three or more), sex, and citizenship status. Note that doctrinal changes are absorbed in the time dummies. All standard errors are heteroskedasticity robust. Sample is all serious crimes from Guidelines years 1992–2009. N=524,315.

Table 2: Three-Year Window around *Koon*

	(1) Departure (Linear Probability)	(2) Departure (Odds Ratio)	(3) Log Prison	(4) Offense Level	(5) Binding Minimum
Black Defendant	-0.001 (0.004)	0.94 (0.071)	0.25** (0.018)	2.59** (0.28)	-0.012* (0.003)
Hispanic Defendant	-0.024** (0.006)	0.81* (0.072)	0.19** (0.021)	1.56** (0.29)	0.010+ (0.054)
Deferential	-0.004 (0.005)	0.93 (0.062)	-0.012 (0.020)	-0.26 (0.16)	-0.002 (0.007)
Deferential × Black	0.011+ (0.006)	1.21** (0.10)	-0.056* (0.022)	0.11 (0.21)	-0.002 (0.004)
Deferential × Hispanic	0.057** (0.013)	1.66** (0.15)	-0.078** (0.022)	-0.19 (0.30)	0.002 (0.070)
N=67,695					

**significant<.01; *significant at <.05; +significant<.10. Regressions include criminal history controls, base offense level dummies, an interaction term between criminal history and base offense level, district and offense dummies as well as controls for age, age-squared, number of dependents (dummies for one, two, and three or more), sex, and citizenship status. The offense level regressions do not control for base offense level. Sample includes only serious crimes. Robust standard errors are clustered at the district level.

Table 3: Three-Year Window around the PROTECT Act

	(1) Departure (Linear Probability)	(2) Departure (Odds Ratio)	(3) Log Prison	(4) Offense Level	(5) Binding Minimum
Black Defendant	-0.007 (0.004)	0.96 (0.052)	0.13** (0.022)	2.53** (0.27)	0.010** (0.003)
Hispanic Defendant	0.007 (0.006)	0.99 (0.037)	0.092** (0.029)	1.57** (0.29)	0.025** (0.006)
Deferential	0.034** (0.004)	1.56** (0.086)	-0.062** (0.020)	-0.26 (0.16)	-0.002 (0.003)
Deferential × Black	-0.015** (0.005)	0.90 (0.074)	0.034 (0.024)	0.11 (0.21)	-0.002 (0.004)
Deferential × Hispanic	0.009 (0.013)	0.97 (0.12)	-0.007 (0.034)	-0.19 (0.30)	-0.007+ (0.004)
N=101,946					

**significant<.01; *significant at <.05; +significant<.10. Regressions include criminal history controls, base offense level dummies, an interaction term between criminal history and base offense level, district and offense dummies as well as controls for age, age-squared, number of dependents (dummies for one, two, and three or more), sex, and citizenship status. The offense level regressions do not control for base offense level. Sample includes only serious crimes. Robust standard errors are clustered at the district level.

Table 4: Three-Year Window around *Booker*

	(1) Departure (Linear Probability)	(2) Departure (Odds Ratio)	(3) Log Prison	(4) Offense Level	(5) Binding Minimum
Black Defendant	-0.021** (0.007)	0.83** (0.051)	0.14** (0.020)	0.093 (0.069)	0.012** (0.004)
Hispanic Defendant	-0.034** (0.006)	1.09 (0.068)	0.097** (0.024)	-0.24** (0.061)	0.029** (0.008)
Post-Booker	0.14** (0.008)	2.86** (0.22)	-0.036* (0.016)	0.18** (0.046)	0.019** (0.006)
Post-Booker × Black	0.005 (0.009)	1.25* (0.11)	0.000 (0.019)	-0.11 (0.067)	0.001 (0.004)
Post-Booker × Hispanic	0.005 (0.012)	0.84* (0.06)	0.010 (0.024)	0.037 (0.090)	-0.007 (0.004)
N=107,633					

**significant<.01; *significant at <.05; +significant<.10. Regressions include criminal history controls, base offense level dummies, an interaction term between criminal history and base offense level, district and offense dummies as well as controls for age, age-squared, number of dependents (dummies for one, two, and three or more), sex, and citizenship status. The offense level regressions do not control for base offense level. Sample includes only serious crimes. Robust standard errors are clustered at the district level.

Table 5: Three-Year Window around *Rita, Gall, and Kimbrough*

	(1) Departure (Linear Prob.)	(2) Departure (Odds Ratio)	(3) Log Prison	(4) Offense Level	(5) Binding Minimum	(6) Binding Minimum (>10 years)
Black Defendant	-0.027** (0.006)	0.99** (0.030)	0.11** (0.021)	1.08** (0.14)	0.011* (0.005)	0.090* (0.004)
Hispanic Def.	-0.043** (0.008)	1.03** (0.003)	0.10** (0.023)	0.85** (0.21)	0.024** (0.005)	0.014** (0.003)
Post-RGK	0.071** (0.006)	1.52** (0.058)	-0.095** (0.014)	-0.20+ (0.11)	0.009** (0.003)	0.008* (0.003)
Post-RGK × Black	-0.019** (0.007)	0.92+ (0.041)	0.12** (0.021)	-0.17 (0.12)	0.029** (0.006)	0.020** (0.004)
Post-RGK × Hispanic	0.004 (0.009)	1.03 (0.06)	0.035 (0.023)	0.17 (0.19)	-0.006 (0.007)	0.005 (0.006)
N=106,854						

**significant<.01; *significant at <.05; +significant<.10. Regressions include criminal history controls, base offense level dummies, an interaction term between criminal history and base offense level, district and offense dummies as well as controls for age, age-squared, number of dependents (dummies for one, two, and three or more), sex, and citizenship status. The offense level regressions do not control for base offense level. The six-month period between *Rita* and *Gall/Kimbrough* is excluded. Sample includes only serious crimes. Robust standard errors are clustered at the district level.

Table 6: Three-Year Window around *Rita, Gall, and Kimbrough*, by Binding Minimum Category

	Criminal History Category 1/No Firearms Offenses (Mandatory Minimums Less Binding)				Criminal History Category >1 & Firearms Offenses (Mandatory Minimums More Binding)			
	(1) Binding Minimum	(2) Binding Minimum	(3) Log Prison	(4) Log Prison	(5) Binding Minimum	(6) Binding Minimum	(7) Log Prison	(8) Log Prison
Black Def.	0.026** (0.006)	0.026** (0.006)	0.31** (0.036)	0.11** (0.037)	0.018** (0.006)	0.008 (0.006)	0.19** (0.024)	0.11** (0.026)
Hispanic Def.	0.030** (0.006)	0.028** (0.006)	0.35** (0.095)	0.12** (0.040)	0.004 (0.004)	0.020** (0.006)	0.11** (0.032)	0.11** (0.026)
Post-RGK	0.011* (0.004)	0.012** (0.004)	-0.16** (0.039)	-0.13** (0.028)	0.004 (0.006)	0.005 (0.004)	-0.095** (0.021)	-0.11** (0.020)
Post-RGK × Black	0.008 (0.010)	0.013 (0.010)	0.024 (0.052)	0.084* (0.039)	0.034** (0.006)	0.034** (0.006)	0.074** (0.024)	0.10** (0.027)
Post-RGK × Hispanic	-0.012 (0.008)	-0.012 (0.008)	0.065 (0.060)	0.051 (0.040)	0.003 (0.008)	0.001 (0.008)	0.075* (0.031)	0.039 (0.030)
Base controls	No	Yes	No	Yes	No	Yes	No	Yes
Sample Size	42,534				63,696			

**significant<.01; *significant at <.05; +significant<.10. Regressions include criminal history controls, base offense level dummies, an interaction term between criminal history and base offense level, district and offense dummies as well as controls for age, age-squared, number of dependents (dummies for one, two, and three or more), sex, and citizenship status. The six-month period between *Rita* and *Gall/Kimbrough* is excluded. Sample includes only serious crimes. Robust standard errors are clustered at the district level.

Appendix Figure 1: United States Sentencing Guidelines Grid
 Sentencing Table (in months of imprisonment)

Offense Level	Criminal History Category (Criminal History Points)					
	I (0 or 1)	II (2 or 3)	III (4, 5, 6)	IV (7, 8, 9)	V (10, 11, 12)	VI (13 or more)
Zone A	1	0-6	0-6	0-6	0-6	0-6
	2	0-6	0-6	0-6	0-6	1-7
	3	0-6	0-6	0-6	0-6	2-8
	4	0-6	0-6	0-6	2-8	4-10
	5	0-6	0-6	1-7	4-10	6-12
	6	0-6	1-7	2-8	6-12	9-15
	7	0-6	2-8	4-10	8-14	12-18
	8	0-6	4-10	6-12	10-16	15-21
Zone B	9	4-10	6-12	8-14	12-18	18-24
	10	6-12	8-14	10-16	15-21	21-27
Zone C	11	8-14	10-16	12-18	18-24	24-30
	12	10-16	12-18	15-21	21-27	27-33
Zone D	13	12-18	15-21	18-24	24-30	30-37
	14	15-21	18-24	21-27	27-33	33-41
	15	18-24	21-27	24-30	30-37	37-46
	16	21-27	24-30	27-33	33-41	41-51
	17	24-30	27-33	30-37	37-46	46-57
	18	27-33	30-37	33-41	41-51	51-63
	19	30-37	33-41	37-46	46-57	57-71
	20	33-41	37-46	41-51	51-63	63-78
	21	37-46	41-51	46-57	57-71	70-87
	22	41-51	46-57	51-63	63-78	77-96
	23	46-57	51-63	57-71	70-87	84-105
	24	51-63	57-71	63-78	77-96	92-115
	25	57-71	63-78	70-87	84-105	100-125
	26	63-78	70-87	78-97	92-115	110-137
	27	70-87	78-97	87-108	100-125	120-150
	28	78-97	87-108	97-121	110-137	130-162
	29	87-108	97-121	108-135	121-151	140-175
	30	97-121	108-135	121-151	135-168	151-188
	31	108-135	121-151	135-168	151-188	168-210
	32	121-151	135-168	151-188	168-210	188-235
	33	135-168	151-188	168-210	188-235	210-262
	34	151-188	168-210	188-235	210-262	235-293
	35	168-210	188-235	210-262	235-293	262-327
	36	188-235	210-262	235-293	262-327	292-365
	37	210-262	235-293	262-327	292-365	324-405
	38	235-293	262-327	292-365	324-405	360-life
	39	262-327	292-365	324-405	360-life	360-life
	40	292-365	324-405	360-life	360-life	360-life
	41	324-405	360-life	360-life	360-life	360-life
	42	360-life	360-life	360-life	360-life	360-life
	43	Life	Life	Life	Life	Life

Appendix Table I: Descriptive Statistics (standard deviation in parentheses)

	Full Sample	Serious Crimes
<u>Judicial Variables</u>		
Deferential Review	0.664	0.646
De Novo Review	0.336	0.354
Booker	0.237	0.232
<u>Guidelines Variables</u>		
Total Prison (months)	49.55 (67.5)	88.4 (84.0)
Drug Trafficking	38.1	70.1
Bank Robbery	2.8	5.2
Firearms	8.8	16.1
All other	51.3	8.5
Final Offense Level	17.6 (8.76)	23.2 (8.21)
Base Offense Level	16.4 (9.98)	24.2 (24.2)
Any Imprisonment	0.830	0.940
Downward Departure	0.192	0.131
Upward Departure	0.013	0.014
Criminal History I	0.510	0.442
Criminal History II	0.112	0.120
Criminal History III	0.144	0.156
Criminal History IV	0.086	0.093
Criminal History V	0.052	0.055
Criminal History VI	0.094	0.132
<u>Offender Demographics</u>		
Age	34.5 (10.8)	33.3 (10.0)
Male	0.856	0.910
Black	0.258	0.416
Hispanic	0.389	0.232
White (non-Hispanic)	0.320	0.330
Citizen	0.658	0.81
Sample size	969,320	524,315