Michigan Law
UNIVERSITY OF MICHIGAN LAW SCHOOL

RACIAL DISPARITY IN FEDERAL CRIMINAL CHARGING AND ITS SENTENCING CONSEQUENCES

SONJA B. STARR
AND
M. MARIT REHAVI

Racial Disparity in Federal Criminal Charging and its Sentencing Consequences

M. Marit Rehavi* and Sonja B. Starr**

May 7, 2012

Abstract

Using rich data linking federal cases from arrest through sentencing, we assess the contribution of prosecutors' initial charging decisions to large observed black-white disparities in sentence length. Pre-charge characteristics, including arrest offense and criminal history, can explain about 80% of these disparities, but substantial gaps remain across the distribution. On average, blacks receive almost 10% longer sentences than comparable whites arrested for the same crimes. At least half this gap can be explained by initial charging choices, particularly the filing of charges carrying mandatory minimum sentences. Prosecutors are, ceteris paribus, almost twice as likely to file such charges against blacks.

INTRODUCTION

Black male defendants in U.S. criminal cases receive much longer prison sentences than white men do: for instance, in federal courts, the average sentence during 2008 and 2009 was 55 months for whites and 90 months for blacks (U.S. Sentencing Commission 2010). But does this stem from disparities in the criminal justice system (as opposed to differing crime patterns), and if so, at what point in the criminal process do those disparities emerge? This study investigates those questions by constructing a linked multi-agency dataset that follows federal cases from arrest through to sentencing. Detailed charge and arrest data allow analysis of the crucial role of prosecutors' charging decisions, as well as that of underlying offense differences, in producing sentencing disparities. We find substantial black-white gaps in initial charge severity conditional on arrest offense and other observables, including dramatic disparities in the use of "mandatory minimum" charges. These initial charging decisions can explain at least half, and perhaps substantially more, of the black-white sentencing disparities that are not otherwise explained by pre-charge observables.

Prosecutors are the most powerful decision-makers in American criminal courts. The United States essentially has a system of negotiated justice—in federal courts, guilty plea rates are approximately 97%. Prosecutors enjoy vast discretion to choose initial charges, bargain, and stipulate to the facts on which sentences are based. Yet empirical studies of racial disparities in sentencing overwhelmingly assess judicial decisions in isolation, effectively conditioning on the outcome of this series of prosecutorial choices. This empirical literature tells us little about the underlying disparities in prosecutorial decision-making, especially in initial charging, a decision that sets the starting point for bargaining and usually determines the crime of conviction. This paper accordingly asks:

- Do black defendants face more (or less) severe initial charges than comparable white defendants arrested for the same offenses?

- Do black defendants ultimately face more (or less) severe sentences than comparable white defendants arrested for the same offenses?

- How much of the racial disparity observed in sentences is driven by disparities in initial charges?

Prior empirical research on racial disparities in sentencing has not addressed these questions for at least two reasons. First, the sentencing research has overwhelmingly relied on da-

1

tasets drawn solely from the sentencing stage, which typically do not record arrest and initial charge data. We overcome this obstacle by using linking files provided by the Bureau of Justice Statistics to connect data from arresting agencies to case data from prosecutors, the courts, and the U.S. Sentencing Commission; the rich arrest data provide a substantially better proxy for the underlying criminal conduct than those used in existing research. A second challenge impeding research on charging is the lack of readily available scales for measuring the relative severity of charges. Accordingly, we construct several new charge severity measures based on comprehensive legal coding of the statutes and sentencing guidelines governing every federal crime charged in the study period.

We analyze federal cases from the most recent years available, fiscal years 2007 through 2009, involving black and white U.S. citizens arrested for violent, property/fraud, weapons, and public order offenses. While other categories of crimes (most notably drugs) were excluded from the main analysis due to data limitations discussed below, more limited analyses conducted with broader samples suggest fairly similar racial disparity patterns. The sample is limited to males because racial disparity patterns differ by gender, and those affecting black males are of particular policy importance. Black men in the U.S. are incarcerated at a rate over six times that of white men and over ten times that of black women, and men make up over 80% of federal defendants in the sample period.[1]

Using both regression methods and a semi-parametric reweighting approach, we find that black defendants face significantly more severe charges than whites even after controlling for criminal behavior (arrest offense, multiple-defendant case structure, and criminal history), observed defendant characteristics (e.g., age, education), defense counsel type, district, county economic characteristics, and crime rates.[2] Unexplained racial disparities exist across the charge-severity distribution, especially at the high end. The most striking disparities are found in the use of charges that carry non-zero statutory minimum sentences (herein "mandatory minimums"). Ceteris paribus, black men are, on average, nearly twice as likely as white men to face such a charge. The Discussion considers possible explanations for these disparities, including but not limited to statistical and preference-based discrimination.

---

[1] Gender disparities and race-gender interactions in case processing are assessed in Starr (2012).
[2] Counsel type, criminal history, and education were recorded for only subsets of the main charging sample, but within those subsets results were robust to the addition of those variables.

2

The ultimate importance of these charging disparities rests on whether they persist through the judicial process and affect sentences. They do. There is a large raw black-white sentencing gap across the distribution of sentences, as illustrated in Figure 1. Using Oaxaca-Blinder decomposition methods (Blinder (1973); Oaxaca (1973); Oaxaca & Ransom (1994 and 1999)) in combination with Recentered Influence Function (RIF) regressions at the unconditional deciles of the sentence-length distribution (Firpo, Fortin, & Lemieux (2009)), this gap can be decomposed into components explainable by observed characteristics. To that end, we decompose the total black white disparity into that which can be explained by arrest offense, criminal history, and other pre-charge case attributes and components explainable by prosecutors' initial charging choices. The remaining unexplained disparities may reflect disparities in judicial sentencing decisions and/or in the intervening stages of case processing (plea bargaining and sentencing fact-finding).

This approach shows that most, but not all, of the gap in sentence length can be explained by differences in cases' pre-charging traits, especially arrest offense and criminal history. When decompositions are performed using those variables alone, black sentences are still nearly 10% higher at the mean than white sentences, with significant gaps throughout the distribution with the largest unexplained disparities at the highest deciles. To a large extent, however, these remaining sentence gaps can be explained by prosecutors' initial charging decisions—particularly the choice to bring mandatory minimum charges. The mandatory minimum indicator can explain nearly seven percentage points of the race gap at the mean and has a significant and sizeable explanatory effect at every decile, ranging from five to nine percentage points (which is 10 to 20% of the raw observed racial gap at each decile). Nor does this explanatory power merely trade off with that of other covariates: adding the mandatory minimum indicator reduces the total unexplained disparity by half on average and renders it insignificant at several deciles. Approximately five percentage points on average remain unexplained, and this gap may be attributed to the combined effect of all post-charging stages in the justice process: plea bargaining, sentencing fact-finding, and sentencing.

The importance of mandatory minimum charges in explaining sentencing disparities is particularly striking because this sample has a relatively low prevalence of such charges (13 percent, including 18 percent of black defendants' cases); for many offenses, especially outside the context of drug crimes, mandatory minimums are simply not a prosecutorial option. Yet dispari-

ties in the initial charging of quite a small fraction of cases appear to be a major driver of otherwise-unexplained sentence disparity in the pool as a whole. Moreover, our estimates are likely underestimates of the role of mandatory minimums because of the use of highly conservative coding assumptions when charge data were ambiguous. Robustness checks using alternate coding choices suggest that had it been possible to identify all of the mandatory minimum charges that were brought, these charges could potentially explained virtually all of the sentencing gap.

These findings suggest that existing empirical scholarship and policy debates on racial disparity in the criminal justice system have overlooked arguably the most important actor: the federal prosecutor. Recently, many policymakers, led by the U.S. Sentencing Commission (2010), have attributed the apparent growth in the black-white sentencing gap to a recent expansion in judicial discretion: the Supreme Court's 2005 decision in *United States v. Booker* rendered the previously-mandatory U.S. Sentencing Guidelines merely advisory. But it may be too hasty to blame judges for post-*Booker* racial disparities. *Booker* did not completely free up judges; statutory mandatory minimums are still binding, and there are strong racial disparities in their application by prosecutors. Indeed, we find that a surprisingly large part of the black-white sentencing gap today arises in the very cases where judges have the *least* sentencing discretion—those in which mandatory minimums tie their hands.[3]

I.   Prosecutorial Discretion and Disparity

A. Prosecutorial and judicial discretion

U.S. federal prosecutors (like state prosecutors) possess enormous discretion, beginning with the initial choice of what, if any, charges to bring. The federal criminal code, which was once fairly narrow, is now quite vast.[4] The definitions of crimes in the code often overlap, which gives prosecutors a wide menu of charging options in a given case, and moreover, prosecutors must subjectively assess the strength of evidence and choose how to characterize ambiguous facts. For instance, if a gun is found in the car that transported a defendant to a burglary, the prosecutor must decide whether to allege that the burglary legally qualified as a "crime of vio-

---

[3] This study uses post-*Booker* data only; Rehavi & Starr (2012) evaluates the effect of and responses to the judicial discretion created by *Booker*.

[4] The code covers both distinctive areas of federal interest (such as immigration and national security crimes) and crimes similar to those prosecuted in state courts, such as violent crimes and property offenses, some of which are defined to require some federal jurisdictional basis such as a nexus to interstate commerce or federal land or benefits. Weapons offenses occupy a substantial fraction of the federal docket (and the largest share, about 42%, of our sample), as do fraud cases. Federal crimes are on average more severe than state crimes, but they range from petty offenses carrying fines to those carrying life sentences.

lence" (18 U.S.C § 16), that the gun qualified as a "firearm" and that the defendant "carried" it "during and in relation to" the burglary—all of which are necessary to trigger a 5-year mandatory minimum sentence under 18 U.S.C. § 924(c), which would run consecutively to the burglary sentence. A lenient prosecutor might choose to "swallow the gun" and just charge burglary.

The existing literature suggests that prosecutors do not simply always seek to maximize a defendant's potential sentence. Prosecutorial discretion is essentially the power to be lenient, and prosecutors might choose to be lenient for a variety of reasons. Incentives shaping prosecutorial choices include pressure to maximize convictions or sentences, career concerns, resource constraints, and the desire to reduce workload and risk by encouraging guilty pleas (for example: Baker & Mezzetti (2001), Easterbrook (1983), Glaeser, Kessler & Piehl (2000), Kessler and Piehl (1998), Landes (1971), and Prescott (2006)). In addition, prosecutors may act out of sympathy or a sense of fairness. Schulhofer & Nagel (1997, p. 1290) reviewed federal prosecutors' case files and found that charges were frequently manipulated to avoid sentencing consequences that were perceived as excessively harsh.

The initial charging stage is particularly critical. At that stage, the prosecutor enjoys essentially unilateral legal decision-making authority,[5] whereas subsequent plea deals must be agreed to by defendants (under counsel's advice).[6] Despite the 97% plea rate, the initial charge is usually the final charge,[7] in part because DOJ policy discourages subsequent "charge bargaining."[8] Instead, plea-bargaining often focuses on sentence recommendations and factual stipulations concerning aggravating and mitigating factors that will influence the sentence (e.g., the defendant's leadership role in a conspiracy, an aggravator, or his "acceptance of responsibility," a mitigator). Initial charge is the starting point for bargaining over all these terms.

Legal scholars, judges, and practitioners broadly agree that these prosecutorial decisions play a dominant role in determining sentences (for example, see Stith 2008; Miller 2004; Johnson & Gilbert 1996). This was especially true between 1987 and 2005, when the then-mandatory

---

[5] The initial charges are, of course, the first move in a strategic game and are therefore likely to be chosen into account expectations of judicial behavior and the defendant's bargaining position.

[6] Although felony indictments are issued by grand juries on prosecutorial request, this is a mere formality; prosecutors get the charges they seek in 99.9% of cases (Skolnik 1999).

[7] In our sample, overall charge severity was reduced during plea bargaining in 10 to 15% of cases, depending on the severity metric used; in about 85% of cases, the lead charge was identical at the initial and final stages.

[8] In 2003, Attorney General Ashcroft issued a memo ordering line prosecutors to get special approval for charge reductions (Ashcroft 2003). The Ashcroft memo also ordered that the initial charge be the "most serious" offense that is "readily provable"— if obeyed, this would essentially eliminate charging discretion. But this part of the policy contains no special supervision requirements, and may be unenforceable in practice (Stith 2008; Miller 2004).

U.S. Sentencing Guidelines limited judicial discretion to narrow sentencing ranges based on factors mainly determined by plea agreements (the charges of conviction and the aggravating and mitigating "sentencing facts") and the defendant's prior criminal history.[9] In 2005, the Supreme Court's *Booker* decision rendered the Guidelines merely advisory, potentially somewhat weakening prosecutors' sentencing influence. But judges must still at least consult the Guidelines,[10] and moreover, *Booker* did not affect the binding nature of sentencing *statutes*. While most crimes have broad statutory sentencing ranges and no statutory minimum, some criminal statutes do specify "mandatory minimums," and even post-*Booker*, prosecutors can tie judges' hands by bringing such charges.

Prior empirical studies of racial and other demographic disparities in sentencing have consistently considered judicial sentencing decisions in isolation from the prosecutorial choices that preceded them. For instance, federal sentencing studies typically rely on Sentencing Commission data alone and estimate disparities in the final sentence conditioned on the "presumptive legal sentence," usually defined as the low end of the Guidelines sentencing range, which in turn is determined by the crime of conviction and the sentencing facts. Studies using this approach have produced varied findings, with most finding disparities favoring whites (Mustard (2001) reviews the literature). Recently, the U.S. Sentencing Commission found fairly large black-white disparities that it attributes to the expansion of judicial discretion after *Booker* (Sentencing Commission 2010; see also Ulmer, Light, & Kramer (2012), who find smaller disparities).

Black defendants receive statutory "mandatory minimum" sentences more frequently than whites do, suggesting that these charges may be an important mechanism for disparities (Fischman and Schanzenbach (2011); Sentencing Commission (2011(a)). This raises the crucial question of *why* blacks more frequently face mandatory minimum sentences: differences in underlying crime patterns, different rates of qualifying for sentencing-stage loopholes, or disparate exercise of prosecutorial charging discretion? Prior studies have not disentangled these explanations; doing so requires charge and arrest data.

Without accounting for possible disparities in the earlier stages of processing, estimates of sentencing-stage disparity are likely to be biased. The pools of white and black defendants

---

[9] Findings of sentencing facts are entered by judges, but judges usually defer to factual stipulations in plea agreements (Johnson & Gilbert 1996; Schulhofer & Nagel 1997).

[10] See 18 U.S.C. 3553. The guidelines thus remain a key reference point (Scott (2012); see also Bushway, Owens, and Piehl (2012), who find that state advisory guidelines influence sentences).

with a given "presumptive legal sentence" may have different underlying offense characteristics. Moreover, the negotiating parties and the judge are well aware of the sentencing implications when determining the charges and sentencing facts on which the presumptive sentence is based, and act in anticipation of the sentencing consequences, introducing endogeneity concerns.

The limited existing literature on racial gaps in prosecutorial decision-making conversely does not estimate their sentencing consequences. There are almost no studies of racial disparities in *initial* charge severity; one exception is Miethe (1987), who found significant disparities favoring whites in a small sample of Minnesota cases. No existing studies focus on racial disparity in initial charging of mandatory minimums, but research more generally points to the important role of prosecutorial discretion in the application of mandatory minimums. For instance, Bjerk (2005) finds evidence that prosecutors manipulate charges to avoid triggering state three-strikes laws, an extreme form of mandatory minimum. Finally, there are a few studies of disparities in plea-bargaining; these generally find that prosecutors favor white defendants (see Shermer & Johnson (2010), who review this literature but find no disparity in their own study).

II. Data and Methods

A. Data

This study's data come from four sources: the U.S. Marshals' Service (USMS), the Executive Office of the U.S. Attorneys (EOUSA), the Administrative Office of the U.S. Courts (AOUSC) and the U.S. Sentencing Commission. These databases were then matched using dyadic linking files provided by the Bureau of Justice Statistics. The charging sample consists of male U.S. citizens with USMS arrest records and US Attorney files opened in the 2007-2009 fiscal years (the most recent data available); the sentencing sample is cases sentenced in those years. These years are entirely after the 2005 *Booker* decision. Since *Booker* gave judges the most sentencing discretion they have had in decades, these years should be expected to involve *less* prosecutorial influence over sentencing than earlier years.

The principal sample consists of property and fraud crimes, weapons offenses, regulatory offenses, and violent crimes (Table 1). Immigration cases and other cases involving non-citizens were excluded because the stakes in those cases center on deportation, making them not directly comparable to other crimes. Of necessity, drug and child pornography cases were are also ex-

cluded from the main analysis due to limitations in the initial charge data.[11] However, as discussed below, analysis of the sentencing data suggests generally similar racial disparity patterns when drug and child pornography cases are included in the sample.

Demographic data are drawn from the USMS data, the only source available for the entire sample. The USMS records five race categories: white, black, Asian, Native, and other/unknown; the last three groups are too small to analyze and are excluded. Hispanic ethnicity is not identified by USMS, so the white group includes an unknown but undoubtedly substantial portion of Hispanics, potentially downward-biasing the black-white disparity estimates.[12] The sample is limited to cases from the 50 U.S. states and the District of Columbia. Judicial districts lacking minimal racial diversity (listed in the Data Appendix) were excluded because within-district black-white comparisons would not have been meaningful.

After all exclusions (which are further detailed in the Data Appendix) and after further restricting the sample to cases filed in district court (rather than those resolved by magistrates)[13] and for which all relevant variables are recorded, the sample for the main charging analysis is approximately 36,000 cases, with the exact number varying by outcome measure. The main sentencing-stage sample is further limited to cases in the Sentencing Commission data, namely those sentenced for non-petty offenses; it consists of 32,346 cases.[14]

USMS codes the offense recorded by arresting officers with over 400 arrest offense codes. The main specifications reported below condense these into 134 groups, combining codes used in very few cases with others that are sufficiently related to describe the same legal offense; results are robust to the use of the original codes.[15] The arrest offense represents the closest proxy available for the defendant's actual underlying criminal conduct, albeit an imperfect one, as discussed further below. Together with additional police notes (included in the USMS files for a subset of the cases and used in some specifications), the arrest offense also represents a

---

[11] Drug quantity, a critical fact for offense severity, is not reliably recorded in the charging-stage data for this sample. Similarly, the child pornography charge data does not distinguish "receipt" charges from "possession," and the former carries a high mandatory minimum sentence.

[12] Ethnicity data available for sentenced defendants indicates that about 16% of sentenced white defendants and 1% of blacks are Hispanic. If non-Hispanic whites experience advantages relative to both blacks and Hispanics (*see, e.g.*, Sentencing Commission 2010), conflating Hispanics with non-Hispanic whites would lead to underestimates of black-white disparities.

[13] As discussed below, magistrates and not district courts resolve most misdemeanor cases. We separately assess disparities in rates of district court filing to account for possible selection bias.

[14] The results are robust to the loosening the sample restrictions that it was possible to relax.

[15] Appendix Tables A2, A4 and A6.

strong proxy for the information that is available to the prosecutor at the time of charging. Prosecutors generally receive their initial information about the case from law enforcement.

Other controls drawn from USMS include district, age and marital status. County-level economic variables (per capita income, unemployment rate, and poverty rate) and a variable for the violent crime rate per 1000 county residents were constructed by linking AOUSC county fields to data from the 2000 Census and from the 2007-2009 FBI Uniform Crime Reports.[16] An indicator for whether more than one defendant was charged in the case was constructed based on AOUSC data. This variable is included because multi-defendant cases often involve alleged conspiracies with more serious charges, and there could be racial disparities in the rate of involvement in group criminal enterprises.

Certain fields collected by the Sentencing Commission were available only for sentenced cases, including education and the defendant's criminal history category. AOUSC records offense counsel type, which is a strong proxy for poverty because the government provides counsel if and only if the defendant is indigent. This variable is missing in slightly over half of cases, so it is excluded from the main specification. However, it is included in alternate specifications.

AOUSC records up to five charges at the filing stage and five at the termination stage. However, the AOUSC data do not reliably code the *severity* of these charges beyond the felony-misdemeanor distinction.[17] Instead, charges are simply recorded as the detailed section of the criminal code a defendant is charged with violating (for example, 18 U.S.C. § 924(c)). That is also true of charging-stage datasets generally, and the legal literature has also not yet developed any readily available scale for measuring charge severity—a problem that has been a significant obstacle to research on charging and plea-bargaining. We therefore develop, calculate, and apply several new severity measures to map the initial charges and their combinations into a defendant's sentencing exposure. These measures are based on detailed legal research into the sentencing law governing every federal crime charged during the study period as well as estimates based on actual sentencing practices in a baseline period.[18]

---

[16] The county variables control for possible prosecutorial focus on enforcement in poor or high-crime areas within districts and are rough proxies for the defendant's socioeconomic status.

[17] The AOUSC does list a "severity code" that is purportedly based on the statutory maximum, but it is often misleading, as explained further in the Data Appendix.

[18] The charge severity measures reflect different proxies for the expected sentence associated with the defendant's combination of charges (without considering other characteristics of the defendant or the case). Sentence length provides a common unit of measure and is how the criminal justice system translates its severity judgments into quantified units (e.g., months of prison).

The first measure—the statutory maximum sentence associated with the set of charges—has several theoretical advantages. After *Booker*, the statute provides the only firm legal constraints on sentencing, and the statutory maximum defines the defendant's sentencing exposure. The statutory maximum associated with the initial charges is thus, effectively, the worst possible outcome potentially faced by the defendant going into plea-bargaining. The statutory maximum also determines Congress's formal labeling of the charge's severity, ranging from a Class A felony to a petty offense. The main weakness of this measure is that statutory ranges are often extremely wide, and the maximum is often a large overestimate of the ultimate sentence.

The second measure is based on the Sentencing Guidelines, which, although no longer mandatory after *Booker*, still serve as a reference point for judges. The Guidelines provide much finer-tuned offense-severity distinctions than statutes do, but many of these distinctions are based on sentencing-stage fact-finding: subsequent findings of aggravating and mitigating factors that are not elements of the charges themselves but are subsequently found by judges (usually on the basis of negotiated factual stipulations). Because this study focuses on initial *charge* severity, the Guidelines measure we construct is based only on what the prosecutor charged. We calculate the guideline sentence for each defendant assuming that the statutory elements of all charges are proven, but *no* additional findings of fact are made at sentencing. The principal drawback of this approach is that it misses most of the variability that the aggravating and mitigating factors are meant to capture, lumping together a wide range of conduct. If anything, this imprecision is likely to add noise and bias disparity estimates downward.

The third measure of charge severity is the average sentence associated with the statutory offense in practice—specifically, the mean for white defendants sentenced in 2005 and 2006, the period immediately before the analysis sample and almost entirely after *Booker*. The use of cases completed before the study period ensures that the measure is not a product of the decisions it is being used to assess. Only white defendants' sentences were used so as to avoid biasing the relative ranking of sentences by the differential racial composition of offense categories. This measure is mechanical and requires no complex legal coding or assumptions about sentencing facts. The primary drawback is the need for a sufficiently thick sample of cases for each charge to calculate the mean. Charges sentenced fewer than 30 times among whites in the reference period were excluded, which reduced the sample for the mean-sentence analyses by about 24%.

In addition to these continuous measures, a "mandatory minimum" indicator denotes whether any of the initial charges carried a positive statutory minimum sentence. Most criminal statutes set out sentencing ranges that start at zero—that is, there is no minimum. But when a "mandatory minimum" charge is brought, it can powerfully affect the sentence. After *Booker*, mandatory minimums are the prosecutor's only way to bind the judge to a higher sentence.

The construction of these measures is limited by the precision of the AOUSC's recorded charges, which sometimes encompass multiple sub-provisions with different expected sentences. The Data Appendix provides details on the methods of resolving ambiguities and on how multiple charges were combined. Summary statistics for all variables described here are provided in Table 1 for the charging-stage and sentencing-stage samples, respectively.

III. Results

A. Disparities in Initial Charging

To assess disparities in whether the defendants charged in district court[19] face charges carrying a statutory "mandatory minimum" penalty, we estimate:

$$logit(M_{iad}) = \alpha + X_{iad}\beta + \delta_a + \delta_d$$

where M is an unobserved latent variable determining whether the prosecutor filed a mandatory minimum charge and i, a, and d index the individual, arrest offense and judicial district, respectively. Prosecutors' offices are organized by district, each headed by a U.S. Attorney. The district fixed effects are included capture any differences in enforcement priorities or norms across districts. X includes age; county per capita income, poverty and unemployment rates and violent crime rates per 1000 population; and whether the case has multiple defendants. In all regressions, standard errors are clustered at the offense-district level to allow for correlated errors due to crime patterns or enforcement priorities within districts. The results are robust to clustering on offense or district alone.

The absolute gap in mandatory minimum charging rates is large: 18% for blacks versus 8% for whites (Table 1). The disparities are not substantially diminished by controlling for arrest offense and other covariates (Table 2, Column 1). The odds ratio (Table 2, Column 1) corre-

---

[19] Detailed data on charging are available only for cases included in AOUSC's records--those in which charges were filed in district court before an Article III judge, which are mostly felony charges. In contrast, most misdemeanors are disposed of by magistrates, typically with no prison sentence. Therefore, the analyses presented here are all conditional on prosecutors' initial discretionary decision to file in district court. To account for possible selection bias, we first estimated the probability of that filing. Conditional on the variables observable in the USMS data (arrest offense, district, and age), there are no significant racial disparities in filing (Table A1).

sponds to a predicted probability of mandatory minimum charge of over 15% for blacks compared to a white base probability of 8% (the white sample mean). Ceteris paribus, blacks are almost twice as likely to face a mandatory minimum charge.

Table 2 (Columns 2-4) also contains estimates for disparities in the mean severity of district court charges, using logged versions of the statutory maximum, guidelines, and past mean sentence measures described above. The estimates are from simple ordinary least-squares (OLS) regressions of the form:

$$log(\text{Severity}_{iad}) = \beta_0 + \beta_1\text{Black} + X_{iad}\gamma + \delta_a + \delta_d$$

where i, a, d, and X are as defined in equation 1 above. Overall, the patterns are quite consistent across all three measures. Black defendants receive charges that are about 6% to 9% more severe than white defendants' (Table 2, columns 2-4).

The OLS estimates are quite robust. Each cell in Table 3 represents the black coefficient from a separate regression; the rows correspond to alternate specifications and the columns to the four charging measures. One concern is that the recording of arrest offenses may mask genuine differences in the underlying criminal conduct. In FY2007, the USMS data contain a text field with the officer's written description of the offense (herein "police notes"). These descriptions sometimes include references to secondary criminal activity (the arrest code reflects the primary offense identified by the officer) or to other details that might affect the severity of charges. Within the FY 2007 cases, the results are robust to the inclusion of indicators for whether the police notes mentioned guns,[20] other weapons, drugs,[21] minor victims, police victims, conspiracy, and racketeering (Table 3, rows 2a and 2b).

---

[20] Possible unobserved racial differences in prevalence of weapons are a particular concern because weapons often trigger mandatory minimums or other sentence enhancements. But the police notes demonstrate that while weapons are more likely to be cited by the arresting officer when the arrestee is black (guns: 42% versus 24% for whites; other weapons: 7% versus 2%), this difference is already captured by the arrest codes; in our sample, cases in which the police mention weapons are almost always coded accordingly (97% for whites and blacks).

[21] Although cases with drug arrest codes are excluded from the sample, drugs can sometimes be present in cases in which the primary offense is something else. The arrest codes do not allow such cases to be excluded, and excluding them on the basis of the eventual charges chosen by prosecutors would introduce sample-selection issues; moreover, exclusion would preclude an accurate account of disparities in gun and violent crime cases, because an important subset of such cases in federal courts have a connection to the drug trade. Under federal law, the most serious offense usually completely determines the sentence (as well as our charge severity coding, as explained in the Data Appendix), so secondary criminal conduct should not be expected to drive the results.

To test this expectation, in addition to the police notes analysis in Table 3, Cols. 2a-2b, two other robustness checks further illustrated that the reported charging disparities are not the result of racially disparate prevalence of drugs among arrestees for other crimes. First, the results are robust to inclusion of an indicator for whether prosecutor recorded, in the pre-charge investigation file, that drugs were seized in the arrest. Second, the results are also

A related concern is that the arrest offense could itself be influenced by pre-arrest prose-cutorial involvement in the investigation; this would likely to downward-bias the disparity esti-mates, because some of prosecutors' preferences are already captured in the offense code. The extent of pre-arrest prosecutorial involvement in federal cases varies widely. Row 3 illustrates that when cases involving the *most* significant such involvement—those in which the indictment precedes the arrest—are excluded, all charging disparity estimates increase slightly.

Rows 4 through 7 of Table 3 show that the results are similar when the sample is con-fined to the South (Row 4), and are also robust to the addition of controls for the defendant's ed-ucation and criminal history, fields available only for cases in the sentencing sample (Row 5a contains estimates from the primary specification for the sub-sample of defendants with educa-tion and criminal history recorded, Row 5b adds education, and Row 5c adds criminal history). Within the subsample for which counsel type is recorded, the results are robust to the addition of an indicator for publicly appointed counsel (compare Table 3, Row 6a and Row 6b). Together with the county-level covariates, the education and counsel results suggest that racial disparity in charging is not being driven by socioeconomic status. The estimates are also robust to the use of the original ungrouped arrest codes, the exclusion of cases from each of the two largest arresting agencies, the FBI and ATF,[22] the addition of non-citizens to the sample, the addition of controls for Hispanic ethnicity and marital status.[23]

The average disparity may mask heterogeneity in the effects of race throughout the dis-tribution of the three continuous charge severity measures. In contrast, the reweighting and de-composition methods developed by DiNardo, Fortin and Lemieux (1996, "DFL") enable the analysis of the full charge severity distribution by reweighting each group to match the endow-

---

robust to the substitution of charging variables that exclude components stemming from drug charges, an approach that means that only the disparity in non-drug charges is measured.

[22] We do not use arresting agency fixed effects because they would be highly collinear with the arrest offense; many federal law enforcement agencies target particular types of crime.

[23] While our focus is on *initial* charging, we also evaluated whether charge disparities persisted after plea-bargaining by estimating regressions corresponding to those presented here, except applying our severity measures to the charges of conviction instead (Appendix Table A3). The mandatory minimum disparity rates remain almost as large at the conviction stage; disparities in the other severity measures are somewhat reduced, but remain moderately large and significant on all but the guidelines scale. Note that not all of the sentencing effect of the initial charge is neces-sarily mediated by the final charge of conviction; the initial charge also influences bargaining over all other terms of plea agreements.

ments of the other group and calculating the counterfactual distribution.[24] Another advantage of DFL is that it does not require the specification of a parametric model of the relationship between defendant and case characteristics and charging decisions.

To that end, consider each case as a vector made up of a charge severity (c), a vector of defendant and case attributes (z) and the defendant's race (r), with F(c,z,r) representing their joint distribution. One can then consider the counterfactual density f(w; r_c=white, r_z=black), that is the charge severity distribution one would expect if white defendants had the characteristics observed for blacks but those characteristics still had the same loading factors observed for white defendants.   DFL show that:

$$f(w; r_c = white, r_z = black) = \int f(c|r, r_c = white)\varphi_z(z)dF(z|r_z = white)$$

Where, $\varphi_z(z)$ is a weighting function defined as:

$$\varphi_z(z) = \frac{dF(z|r_z = black)}{dF(z|r_z = white)} = \frac{Pr(black|z)}{Pr(white|z)}\frac{Pr(white)}{Pr(black)}$$

And Pr(black|z) can be estimated by the predicted probabilities from a probit regression of the defendant's race on all the observed characteristics. A probit was used to estimate the probability a defendant was black as a function of binary indicator variables for arrest offense, district, education, and multiple-defendant case structure, as well as the county socioeconomic and crime variables and the defendant's age. Notably, while blacks are a minority of the US population, they are 45 percent of the sample of charged and sentenced defendants in this period, and there is a large amount of overlap in the observed characteristics of black and white federal defendants (Table 1).[25]

Figures 2a-2c display histograms of the observed and reweighted distributions for each of the continuous charge severity measures. The raw racial disparity is evident from the substantial gaps throughout the observed distributions: whites are disproportionately concentrated in the least severe charges, while blacks are disproportionately found at higher charge severities. Reweighting on observed characteristics generally closes most, but not all, of the gap between the

---

[24]Alternately, one could also use a propensity score matching estimator, in lieu of reweighting. However, Busso, Dinardo and McCrary (2011) demonstrate that reweighting estimators perform better than propensity score matching estimators in finite samples with a large amount of overlap, which is the case in these data.

[25] There is a substantial amount of overlap in the black and white samples. The 25th, median and 75th predicted probabilities were 0.24, 0.45, and 0.68 respectively. As is standard in the literature (see, for example, Heckman, Ichimura and Todd (1997), Heckman, Ichimura, Smith and Todd (1998), and Smith and Todd (2005)) we eliminated observations with no or very little common support. To that end, 3% of observations with predicted probabilities above 0.97 and below 0.03 were trimmed. The estimates are also robust to alternate or no trimming.

two groups throughout the distribution. Consistent with the regression estimates, racial dispari-
ties in charging are greatly diminished, but persist after adjusting for the observed characteristics
of the defendants and cases.

B. Sentence Disparity and its Relationship to Charging

In this Section, we use decomposition methods to estimate sentencing disparities conditional
on arrest offense and other pre-charge covariates, representing the aggregate disparity introduced
throughout the post-arrest justice process. We then add charging variables to the decompositions
to estimate the extent to which these otherwise-unexplained sentencing gaps can be explained by
the charging disparities demonstrated above.

1. Disparities in Whether the Defendant Is Sentenced to Incarceration

Not all defendants receive incarceration sentences. First, some defendants are either not
convicted at all or convicted only of a "petty offense."[26] Second, among those convicted of non-
petty offenses, about one-fourth receive non-prison sentences such as probation and/or fines.
Following widespread practice in the sentencing literature (*e.g.*, Ulmer, Light, & Kramer 2011),
the choice between prison or non-prison sentences is treated as a binary process that precedes the
sentence-length inquiry, rather than including the non-prison sentences in the main analysis as
zeros, which would be substantively problematic.[27] Table 4 provides the "black" odds ratios
from logistic regressions for these two threshold processes: being convicted of a non-petty of-
fense (Col. 1) and being incarcerated conditional on receiving such a conviction (Col. 2). There
are no significant racial differences in the probability of conviction. However, black defendants
may be more likely to be sentenced to incarceration. The estimates of racial differences in incar-
ceration are marginally significant (p-value 0.06) and potentially non-trivial in magnitude.[28] If

---

[26] Such offenses, those less than Class A misdemeanors, are rare among cases filed in district court (about 0.6% of
the sample), virtually never result in incarceration and are often missing from the sentencing data, so we exclude
them from the subsequent analyses.

[27] Non-incarceration sentences are not really "zero" sentences, but cannot readily be translated into a "prison equiva-
lent." Prison sentences of any length carry unique stigma and life disruption, so the practical difference between ze-
ro and one month may be much greater than the difference between one month and two. In addition, the size of
fines, length of probation and other penalties may vary widely across non-incarceration sentences making it difficult
to compare their relative severities. Another advantage of treating incarceration as a separate process is that the
primary method used below to analyze the sentence length distribution (RIF decomposition) cannot be conducted in
the vicinity of large spikes in the distribution, and thus could not have been conducted in the neighborhood of the
non-incarceration observations had they been included as zeros or some other small value. Estimates of the RIF
decompositions at higher deciles for a sample that includes those with no incarceration are similar to those reported
here. An alternative method, DFL reweighting, does include the zeros and is reported below as a robustness check.

[28] These analyses are conditional on the same covariates used in the charging stage analyses, plus (in the incarcera-
tion analysis) criminal history and education, fields available for sentenced cases.

there is racial disparity in incarceration probabilities, it could produce downward-bias in the estimates of disparity in prison sentence length among those incarcerated.

2. Disparities in Prison Sentence Length

The primary method used to assess the factors contributing to racial disparities in sentence length is Oaxaca decomposition of the black-white gap in the Recentered Influence Functions ("RIF") calculated at the deciles of the sentence distribution (Firpo, Fortin and Lemieux (2009) and Fortin, Lemieux, and Firpo (2011)). The RIF approach enables the estimation of the marginal effect of race at selected quantiles of the *unconditional* distribution. The RIF for each decile of the distribution of log sentence-length (excluding non-incarceration sentences) is defined as:

$$\text{RIF}(S; Q_t) = Q_t + \frac{t - I[S \leq Q_t]}{f_s(Q_t)}$$

where Q_t is the sample quantile and $f_s(Q_t)$ is the density at that point. In practice, both the sample quantile and the density are estimated separately for black and white defendants using their respective sentencing distributions.

Coefficients estimated in regressions with the RIF of log sentence length as a dependent variable approximate unconditional quantile estimates at the quantile for which the RIF was calculated.[29] This enables us to directly estimate the racial gap at each chosen decile and then apply traditional decomposition methods. To that end, the gap in the RIF at each decile (as well as the gap in the mean) is decomposed using pooled Oaxaca decompositions (Jann 2008). Unlike the DFL method used for the charging analysis above, the RIF decomposition can jointly numerically quantify the amount of the disparity that could be explained by each of the factors of interest.[30] This is also a vast improvement on traditional quantile regressions, which only produce estimates of the marginal effect of a variable at the qth quantile of the *conditional* distribution.[31]

To that end, for each decile of the distribution of log sentence length we estimate:

---

[29] Following Firpo, Fortin and Lemieux's (2009) terminology, "unconditional quantiles are the quantiles of the marginal distribution of the outcome variable Y, i.e. the distribution obtained by integrating the conditional distribution of Y given X over the distribution of X" (p. 2)

[30] Despite the advantages of RIF, it is not appropriate for the earlier charging analysis (for which DFL was used instead) because the distribution of charge severity is characterized by large spikes, which preclude RIF from consistently approximating the unconditional quantile estimates at those points in the distribution. The distribution of actual sentences, in contrast, is quite smooth once non-incarceration sentences (which would otherwise be a large spike at zero) are removed.

[31] One cannot use the law of iterated expectations to recover unconditional population quantile estimates from the conditional quantile estimates, a key limitation for both policy analyses and Oaxaca-style decompositions.

$$RIF(S_b) - RIF(S_w) = (\overline{X}_b' - \overline{X}_w')\widehat{\beta}^p + [\overline{X}_b'(\widehat{\beta}^b - \widehat{\beta}^p) - \overline{X}_w'(\widehat{\beta}^w - \widehat{\beta}^p)]$$

where b and w stand for black and white respectively and $RIF(S_r)$ is calculated as described above. $\widehat{\beta}^p$ is the vector of coefficients from the pooled OLS regression of the RIF on the observed characteristics and a race indicator variable. The observed characteristics (X) include district and arrest offense fixed effects, defendant characteristics (criminal history,[32] age, education), whether the case involved multiple defendants and characteristics of the county where the defendant was arrested (log per capita income, the crime rate, and the unemployment and poverty rates. Table 5 contains the resulting estimates for the decompositions of the RIF at the odd deciles of the sentence length distribution. The traditional Blinder-Oaxaca decomposition of the mean disparity is displayed in the first column. The first panel (5a) represents the decomposition *without* charging controls. This panel assesses the extent to which the large raw black-white sentencing disparities shown in Figure 1 can be explained by the arrest offense and other pre-charge characteristics of the offender and case.

The decompositions in panel 5a find large raw racial disparities that are mostly, but not entirely, explained by characteristics fixed prior to charging: the arrest offense, criminal history, age, district, number of defendants, education, and county economic and crime measures (Figure 2a). Of these, the largest contributors by far are arrest offense and criminal history. The unexplained racial difference at the mean is roughly 9.6 percentage points, over one-fifth of the total black-white difference in sentences. Thus, even after accounting for observed differences in criminal history and behavior, black defendants are receiving sentences that are nearly 10 percent higher than white defendants. Figure 3a illustrates the total disparity and the "explained" component at each decile; the difference between the bars represents the remaining unexplained racial gaps in sentencing. Significant unexplained differences favoring whites remain in every decile but the first.[33] The unexplained gaps are quite large at the top of the sentencing distribu-

---

[32] Appendix Table A1 illustrates the large effect of criminal history at sentencing. In contrast, criminal history plays no legal role at charging, and Table 3 (rows 5a and 5c) showed that its effect on charging disparity estimates was fairly small.

[33] It is possible that the lack of disparity at the first decile could be the result could be the result of downward bias due to sample selection on the incarceration margin. While only marginally significant (p-value 0.06), estimates of the probability of being incarcerated (Table 4) suggest blacks may be more likely to face incarceration sentences and thus to be in the sample for the RIF analysis. If that is the case, black defendants appearing in the RIF sample may have slightly weaker cases against them, conditional on arrest offense and other covariates, than white defendants do. Sample selection as a result of this disparity is most likely to appear at the bottom of the sentence-length distri-

tion. At the 90th percentile almost half the total observed black-white sentence disparity (20 per-centage points) remains unexplained.

The decompositions in panels 5b and 5c evaluate whether these residual disparities—those unexplained by pre-charge factors—can be explained by prosecutors' charging choices, rather than by the post-charge justice process. Panel 5b shows the results of decompositions identical to those in 5a, except that the mandatory minimum charge indicator is added. The total and "ex-plained" gaps at each decile are displayed in Figure 3b. In Panel 5c, both the mandatory mini-mum indicator and the statutory maximum measure are added; together, these variables reflect the legal constraints on sentencing implied by the initial charge.[34]

As Panels 5b and 5c and Figure 3b illustrate, a large part of the disparities that were not already unexplained by pre-charge factors can be explained by prosecutors' charging decisions. The charging variables are statistically and economically significant explanatory factors at every decile and at the mean. The statutory maximum sentence and the mandatory minimum indicator are capable of explaining about 6.2 percentage points of the black-white gap (which is almost 14% of the *overall* black-white sentencing disparity) at the mean (Table 5c).   Only about 4.8 percentage points of disparity (11% of the overall disparity) remain unexplained at the mean--half the size of the disparity that had been unexplained before the addition of the charging varia-bles. Both charging measures have significant explanatory power at every decile. In all, after the addition of the charging variables, the unexplained gaps favoring whites are substantially re-duced or completely eliminated at all but one decile.

Perhaps most remarkably, the estimates are nearly identical when the mandatory mini-mum indicator is the *only* measure of charge severity.[35] All else equal, the decision to initially charge a defendant under a statute carrying a mandatory minimum alone is capable of explaining approximately five to nine percentage points of the black-white sentencing gap at each decile, and nearly seven percentage points at the mean. Although a portion of that effect appears to

---

bution, because the affected cases are likely to be relatively non-serious, with facts placing them near the borderline between incarceration and nonincarceration sentences.

[34] Alternative specifications substituted the other continuous measures for the statutory maximum. Substituting the mean sentence measure did not alter the analysis substantially, but the Guidelines measure has less (albeit still sig-nificant) capability of explaining sentencing disparities.

[35] Both the estimates of the amount explained by the use of mandatory minimums and the persisting unexplained disparity are similar when the other charge-severity measures (the guideline and mean sentence) are used instead of the statutory maximum.   The amount of the disparity that can be explained by charge severity is greater with the alternate measures. However, as the unexplained is virtually unchanged, this additional explanatory power trades off with that of other variables already in the decomposition.

trade off with the explanatory value of pre-charging variables already in the decomposition (such as the arrest offense), most of it does not. The addition of the mandatory minimum indicator to the decomposition reduces the average unexplained disparity by half, from 9.6 to 4.8 percentage points (compare table 5a to table 5b), and leaves the same unexplained disparity as the decomposition that included both charging variables (table 5c). This suggests that the large disparities documented in the use of mandatory minimum charges exert a particularly powerful influence at the sentencing stage (Figure 3b; Table 5b).

These estimates are very robust to variations in the specification and sample. Table 6 contains alternative RIF decomposition estimates of the amount of racial disparity that can be explained by prosecutors' use of mandatory minimum charges. All use the base specification from Table 5b and accordingly include the mandatory minimum variable but no other charging measures.[36] In all specifications and subsamples, the mandatory minimum indicator is economically and statistically significant at every decile, and generally has a similar pattern of estimates across the deciles. The estimates are robust to excluding prearrest indictment cases (row 2) and limiting the sample to the South (row 3). One might be concerned that prosecutors treat crimes (particularly those involving guns) more harshly in high crime areas and that blacks may be more likely to live in those areas. However, the estimates are similar when the sample is restricted to high crime counties (row 4). The results are also robust to the addition to the counsel type variable (compare rows 5 and 6), a proxy for poverty, and indeed, having publicly appointed counsel does not explain any of the racial gap in sentencing at any decile within the sub-sample for which counsel type is recorded. As in charging, the results are unchanged by the addition of indicators for police notes mentioning guns, other weapons, drugs, child and police victims, conspiracy and racketeering within the pool for which this field is available (compare rows 7 and 8).

The final 3 rows to Table 6 explore whether the mandatory minimum results could be biased by our coding of the mandatory minimum variable, which we coded as 0 when the AOUSC statute information was ambiguous. The extent of the resulting measurement error and possible bias cannot be directly tested at the charging stage, but *can* be tested at the *conviction* stage, where our coding can be compared to the actual mandatory minimum recorded by the sentencing judge (which is in the Sentencing Commission data). At the conviction stage, our coding labels

---

[36] Robustness checks on the versions of the decompositions that include no charge controls and the alternate charge severity measures are available on request.

9% of cases as "mandatory minimums," while the Sentencing Commission data reveal a rate of 16%. Rows 9 and 10 substitute our and the Sentencing Commission's indicators of a mandatory minimum *conviction*, respectively, for the mandatory minimum *charging* indicator. With our coding, mandatory minimum convictions can explain about six percentage points of the mean black-white sentencing gap, similar to the share explained by mandatory minimum charging in the main specification (Table 5b). In contrast, when the Sentencing Commission's coding is used (row 10), the mandatory minimum conviction variable explains about 14 percentage points of disparity on average, and no significant unexplained disparity remains. This suggests that our conservative coding of the presence of a mandatory minimum charge leads us to substantially underestimate mandatory minimums' role in explaining sentencing disparity.[37] In fact, estimates using the Sentencing Commission's coding suggest mandatory minimum convictions could explain *all* of the otherwise-unexplained racial gaps in sentencing in our sample, on average and throughout all but the top of the sentencing distribution.

The use of the Sentencing Commission's mandatory minimum conviction indicator also allows us to consider whether similar disparity patterns might also be found in drug and child pornography cases—categories excluded from the main sample because the AOUSC charge data were inadequate. At least in drug cases (by far the larger of the two categories), there is reason to believe they might be: drug cases have large racial disparities in sentence outcomes (e.g., Sentencing Commission 2010) and clustering of sentences around common mandatory minimum thresholds (Fischman & Schanzenbach, 2011). And given the much higher prevalence of mandatory minimums in both drug and child pornography cases, even smaller race gaps in rates of mandatory minimum charging could have a large impact on overall sentencing disparities.[38] We therefore re-estimate the decomposition with the sample expanded to include drug and child pornography cases (row 11). The results in the broader sample are fairly similar to those in the main sample, with the mandatory minimum conviction indicator explaining over 11 percentage points

---

[37] The differences between rows 9 and 10 (Table 6) are particularly large at the upper end of the sentencing distribution. This is likely because certain particularly long mandatory minimums could not be identified based on the ambiguous AOUSC charge data, which for example do not distinguish (within charges brought under 18 U.S.C. § 1111) between first-degree murder (carrying a mandatory life sentence) and second-degree murder (which has no minimum and a zero-to-life range). Very long mandatory minimums are often only triggered based on special circumstances laid out within statutes, and while we were sometimes able to identify those circumstances based on other charges in the case, often these cases were impossible to identify.

[38] In the sample used in Table 6, Row 11, 58% of whites and 72% of blacks convicted of drug crimes were convicted under statutory provisions that carry mandatory minimum sentences, according to the Sentencing Commission.

of disparity on average—which (although less than the 14 percentage points observed in the narrower sample, row 10) is enough to render the remaining unexplained disparity insignificant. This estimate is suggestive that disparities in mandatory minimum charging are an important contributor to racial gaps in sentencing in these cases as well.[39] These results are only exploratory, however: the lack of drug quantity information means that in drug cases, the arrest offense is a weaker proxy for the severity of underlying criminal behavior.[40]

Like the charging severity results, the sentencing results are also robust to the use of the original arrest, the addition of non-citizens to the sample, the addition of controls for Hispanic ethnicity and marital status, and the exclusion of cases from the FBI and ATF, respectively.[41]

Finally, despite allowing for heterogeneity in the relationships between explanatory variables and the racial sentencing disparities across the distribution, the RIF-Oaxaca decomposition still relies on parametric estimates of the relationship between the observed characteristics and sentence length (the initial pooled regression). It is accordingly susceptible to the limitations of traditional Blinder-Oaxaca decompositions, particularly misspecification and functional form assumptions (see Barsky, Bound, Charles and Lupton (2002) for a discussion of these issues and the advantages of reweighting methods).[42] Figure 4 illustrates the robustness of the results to an alternate decomposition method that does not require the assumptions of the RIF/Oaxaca approach. The figure contains the results of a DFL reweighting exercise analogous to the one per-

---

[39] The disparities in the pool that includes drugs are, however, substantially smaller at the high end of the distribution. This is probably because the mandatory minimum variable is only an indicator for whether a conviction offense carries any level of mandatory minimum, an approach that is less effective in capturing variation in a pool in which mandatory minimums are very common and the *length* of mandatory minimum is an important source of disparity. Even *white* drug defendants face some mandatory minimum in about 58% of cases in the Row 11 sample, so the mandatory minimum indicator will necessarily not explain much racial disparity in the upper half of the distribution of drug cases (the effects seen at the high end in Row 11 are probably coming from the non-drug part of the pool). When the Row 11 analysis is repeated substituting a three-level categorical variable that differentiates among lengths of mandatory minimum recorded by the Sentencing Commission (three levels are enough to capture the main variation in drug minimums, which are generally zero, five, or ten years), that variable explains about fourteen percentage points of disparity on average, including larger disparities in the upper end of the distribution. These results closely track the patterns in Row 10 for the main sample. In contrast, substituting a categorical mandatory minimum variable for the indicator has little effect on the results for the main sample.

[40] However, there is reason to believe that controlling for drug quantity at arrest, if it were possible, might not change the results. The EOUSA investigation file records the type and quantity of drugs seized at arrest; unfortunately, there are serious defects in the quantity field beginning in 2004. But for the most recent three years in which quantity information is available, FY 2001-2003, when drug quantities are converted to their "marijuana equivalents" under sentencing law (a method allowing comparison across drug types), there is no racial disparity in the average amount seized, conditional on arrest offense and other pre-arrest covariates.

[41] Appendix Tables A4 and A6.

[42] Given the large overlap in observed characteristics in the black and white samples, Barsky et al's (2002) concerns about bias introduced by extrapolating outside the observed data ranges is of less concern in this setting. However, their larger point about the potential for bias due to misspecification and functional form assumptions remain.

formed in the charging analysis (Figs. 2a-2c), but with weights calculated using all of the covariates included in the RIF sentencing decompositions, including the mandatory minimum charging indicator. The reweighting sample includes the non-incarceration sentences, thus also relaxing the sample selection imposed above.[43] The results of the reweighting exercise are consistent with the decompositions, indicating that nearly all of the observed racial disparity in sentences can be explained once the groups are reweighted to account for observed differences in defendant and case characteristics and the use of mandatory minimum charges.

IV. Discussion

Conditional on the arrest offense and other observed variables, black arrestees appear to be charged more harshly by prosecutors, especially with respect to mandatory minimum charges. Moreover, these charging disparities translate into substantial sentencing disparities. The findings concerning mandatory minimums' sentencing consequences are particularly striking given that such charges are brought in only 13% of cases in this sample (and only 19% even of black defendants' cases). It is notable that disparities within a type of charge that does not apply to most cases could explain such a large fraction of the otherwise-unexplained disparities in the total pool. The interpretation of these results is not entirely straightforward, however—here, we consider some competing explanations.

First, while the arrest offense is the best proxy available in the data for the defendant's actual criminal conduct, it is naturally an imperfect proxy. Between the 430 arrest codes, the multi-defendant case variable, and the written police notes, the information on arrest is fairly rich, and moving from slightly less detailed arrest information (our grouped codes) to a greater level of detail (by adding the police notes indicator variables, and by using the original ungrouped arrest codes) does not reduce disparity estimates. Moreover, the results do not seem to be driven by any particular arresting agency's cases. But there could still be factual differences that are not captured by the arrest codes or by the written description—like prosecutors, officers could choose to describe the same facts in different ways, and if they do so along racial lines, that divergence could bias racial disparity estimates that are conditioned on the arrest data.[44]

---

[43] DFL does not require any specification of the parametric relationship between underlying characteristics and outcomes, therefore characterization of non-incarceration sentences as zeros is of less concern here.

[44] For example, if police are sympathetic to black arrestees (or believe prosecutors and judges treat them too harshly) and soften their arrest reports accordingly, then prosecutors could appear to charge black defendants more harshly conditional on arrest offense but be charging them equally conditional on true criminal behavior.

However, scholarship on law enforcement suggests that the direction of any bias introduced by arrest-stage discretion is likely downward. Studies in numerous contexts have pointed to the possibility of police bias against minority suspects (e.g., Gelman, Kiss, & Fagan (2007)), although there is considerable debate as to whether such gaps are driven by statistical or preference-based discrimination (e.g., Knowles, Persico, & Todd 2001; Antonovics & Knight 2004). If federal agents are harsher on black suspects, one would expect them to record, on average, more offenses for blacks relative to their true conduct and to up-code offenses whenever possible. Then, prosecutors' cases against black defendants should, conditional on arrest offense, be weaker and the estimated results may understate the "true" race gaps in charging (or at least are unlikely to overstate them).[45]

Another plausible theory is that race is correlated with unobserved characteristics of the *defendant* (rather than the case) that influence prosecutorial choices. Criminal history is one possibility, but it does not appear to explain much charging disparity. Other candidates include poverty and, relatedly, defense counsel quality. But the inclusion of counsel type and other socioeconomic controls (education, marital status, and county-level measures) do not reduce the racial gaps in charging or sentencing. This is less surprising than it appears, given the high quality of federal public defenders (see Posner & Yoon 2011).[46]

While other unobserved differences cannot be ruled out, there remains the plausible possibility that the observed disparities are driven by discrimination, which could be either statistical or preference-based. The mechanism for preference-based discrimination might well be unconscious racial disparities in empathy that drives selective leniency, rather than animus driving selective harshness. (See Fong and Luttmer (2009) for a discussion of race and charitable giving, Linder (1996) for an argument about jurors and racially selective empathy, and Goette, Huffman and Meier (2006) for experimental evidence of own-group bias in third party punishment games.) Statistical discrimination might, for instance, be based on expectations concerning criminal recidivism (see Curry and Klumpp 2009). Prosecutors might also charge in the shadow of expectations of harsher treatment of blacks by judges or juries (*e.g.*, Easterbrook 2003).

---

[45] While the conclusions of the policing studies vary, they at least do not generally suggest discrimination *favoring* blacks. One cannot rule out the possibility that the reverse pattern holds within the sample; federal agents may differ from the state and local police generally studied in the policing literature. But the results at the charging and sentencing stages do not, at least, appear driven by any particular enforcement agencies' patterns. See Appendix Tables A2, A4 and A6.

[46] In contrast, in state courts that have less effective public defender programs, any racial disparities might be expected to be compounded by socioeconomic disparities.

Preference-based and statistical discrimination mechanisms cannot be disentangled using these data. Notably, however, if any form of purposeful race-based discrimination is involved, none of these mechanisms are legally permissible. Otherwise-unconstitutional discrimination cannot be legally defended on the basis of statistical generalizations about group traits, regardless of their empirical support (*J.E.B. v. Alabama ex rel T.B.*, 511 U.S. 127 (1994)). In addition, statistically discriminating prosecutors may not have the opportunity to modify charges as they update their beliefs about the individual or case over time, as employers can do in the labor market (Altonji and Pierret 2001). Criminal cases are often processed quickly, with little chance for personal interaction between prosecutors and defendants, and because DOJ policy discourages charge-bargaining, it may be costly or difficult for prosecutors to change charges even if they do update their beliefs.

CONCLUSION

This study provides robust evidence that black male federal defendants receive longer sentences than whites arrested for the same offenses and with the same prior records. On average black males receive sentences that are approximately 10% longer than comparable white males with those at the top of the sentencing distribution facing even larger disparities. Much of that disparity appears to be driven by decisions at the initial charging stage, especially by prosecutors' filing of "mandatory minimum" charges, which, *ceteris paribus*, they do twice as often against black defendants. Our estimates of disparities in prosecutorial decisions are likely conservative, because they do not encompass gaps introduced by *prearrest* prosecutorial involvement in the case, nor do they account for possible disparities in law enforcement.

The importance of mandatory minimums in sentencing disparity is particularly striking given that our sample consists of crime categories in which mandatory minimums are relatively uncommon and our conservative coding decisions almost certainly led us to underestimate their role quite substantially. Estimates using the Sentencing Commission's recording of the presence of a mandatory minimum at conviction suggest that prosecutors' decisions regarding mandatory minimums could even potentially explain *all* of the otherwise-unexplained racial gaps in sentencing in our sample, at all but the highest deciles. Furthermore, prosecutor' decision-making does not end with the initial charges; they continue to be involved in the plea bargaining and sentencing phases of cases (including negotiating the stipulations that play a key role in sentencing fact-finding), and could also play a role in the remaining unexplained disparity that we attribute to

post-charging stages of the justice process. For these reasons, our estimates should be viewed as a lower bound on the impact of prosecutorial decision-making on sentencing disparity.

Recent policy and scholarly debates about post-*Booker* racial disparities have focused heavily on the disparity risks associated with judicial discretion. This study suggests that those concerns may to a substantial extent be misplaced. Although this study does not attempt to isolate the impact of judicial sentencing decisions (which are only one part of the post-charge process), the results suggest that they are probably only a modestly important source of disparity. Sentence disparities at the mean and at most deciles can be largely explained by three factors: the original arrest offense, the defendant's criminal history, and the prosecutor's initial choice of charges. That leaves less than a 5% mean gap in sentences for the post-charge justice process (including judicial decisions) to explain—perhaps substantially less if our estimates understate the explanatory value of the mandatory minimum, as suggested above. However, there are somewhat larger unexplained gaps at the top of the distribution, so judicial decisions could be producing substantial disparities in the subset of defendants committing the most serious crimes or those with the most extensive criminal history.

But overall, the results suggest that prosecutors' charging decisions are at least as important a source of racial disparity as judicial sentencing decisions are, if not more so. Such prosecutor-driven disparities pose a considerable policy challenge. Substantial prosecutorial discretion is fundamentally ingrained in the U.S. justice system and would be difficult to take away. Even if doing so were desirable, resources do not permit prosecutors to pursue every possible charge in every case, and prosecutors also must assess the strength of evidence, an inescapably discretionary process. Indeed, the racial disparities found here emerged despite the fact that during the entire sample period official DOJ policy purported essentially to eliminate discretion by requiring prosecutors to charge the most serious provable offense (Ashcroft 2003).

Still, even if prosecutorial charging discretion is inevitable, sentencing law can help to shape the amount of influence these decisions have over the sentence. Perhaps the strongest potential policy implications suggested by this study concern mandatory minimum sentences. While racial disparities in the application of mandatory minimums have long been noted, it was previously unclear whether the gap was driven by disparate charging or, instead, by different underlying criminal offense patterns. This study provides evidence that it is not just that sentencing law applies mandatory minimums to crimes disproportionately committed by blacks. Rather,

prosecutors appear to be nearly twice as likely to *use* the laws against black defendants when doing so is a discretionary choice. This suggests that calls by policymakers to respond to post-*Booker* sentencing disparity by expanding mandatory sentencing rules in an attempt to constrain judicial discretion could be counterproductive.

REFERENCES

Altonji, J. G. and C. R. Pierret. "Employer Learning And Statistical Discrimination," *Quarterly Journal of Economics,* 2001, v116(1,Feb), 313-350.

Anderson, J. M., J. R. Kling, & K. Stith. 1999. "Measuring Inter-Judge Sentencing Disparity Before and After the Federal Sentencing Guidelines." *Journal of Law and Economics*, 42: 271-307.

Antonovics, K. & B. G. Knight. 2009. "A New Look at Racial Profiling: Evidence from the Boston Police Department." *Review of Economics and Statistics*, 91: 163-177.

Ashcroft, J. 2003. "Department Policy Concerning Charging Offenses, Disposition of Charges, and Sentencings." *Memorandum*, Sept. 22, 2003.

Baker S. & C. Mezzetti. 2001. "Prosecutorial Resources, Plea Bargaining, and the Decision to Go to Trial," *Journal of Law, Economics and Organization*, 17: 149-167.

Barsky, R., J. Bound, K. Charles, and J. Lupton. 2002. "Accounting for the Black-White Gap: A Nonparametric Approach." *Journal of the American Statistical Association*, 97 (459): 663-673.

Bjerk, D. 2005. "Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion Under Mandatory Minimum Sentencing." *Journal of Law and Economics*, 48: 591-625.

Blinder, A. S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources*, 8: 436-455.

Bushway, S., E. Owens & A. Piehl. 2012. "Sentencing Guidelines and Judicial Discretion: Quasi-experimental Evidence from Human Calculation Errors." *Journal of Empirical Legal Studies*, 9 (2): (in press).

Busso, M., J. Di Nardo and J. McCrary. 2011. "New Evidence on the Finite Sample Properties of Propensity Score ReWeighting and Matching Estimators" *mimeo.*

Curry, P. & T. Klumpp. 2009. "Crime, Punishment, and Prejudice." *Journal of Public Economics,* 93: 73-84.

DiNardo, J., N. M. Fortin, and T. Lemieux. 1996. "Labour Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, 64: 1001-1046.

Easterbrook, F. H. 1983. "Criminal Procedure as a Market System." *Journal of Legal Studies*, 12: 289-332.

Firpo, S., N. M. Fortin, and T. Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica*, 77: 953-973.

Fong, C. M. & E. F. P. Luttmer. 2009. "What Determines Giving to Hurricane Katrina Victims? Experimental Evidence on Racial Group Loyalty". *AEJ: Applied*, 1 (2): 64-87.

Fortin, N. M., T. Lemieux and S. Firpo. 2011. "Decomposition Methods." In O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics,* 4A: 1-102, Amsterdam: North-Holland.

Fischman, J. B. & M. M. Schanzenbach. 2011. "Judicial Discretion and the United States Sentencing Guidelines," Working Paper.

Gilbert, Scott A. & M. T. Johnson. 1996. "The Federal Judicial Center's 1996 Survey of Judicial Experience." *Federal Sentencing Review*, 9: 87-93.

Glaeser, E. L., D. P. Kessler & A. M. Piehl. 2000. "What Do Prosecutors Maximize?." *American Law & Economics Review*, 13 (2): 259-290.

Goette, L., D. Huffman & S. Meier. 2006. "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups." *The American Economic Review*, 96(2): 212-216.

Heckman, J., H. Ichimura, and P. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64(4): 605-654.

Heckman, J., H. Ichimura, J. Smith, and P. Todd. 1998. "Characterizing Selection Bias Using Experimental Data," *Econometrica,* 66(5): 1017-1098.

Jann, B. 2008. "The Blinder-Oaxaca Decomposition for Linear Regression Models." *The Stata Journal*, 8(4): 453-749.

Kessler, D. P. & A. M. Piehl. 1998. "The Role of Discretion in the Criminal Justice System." *Journal of Law, Economics & Organization*, 14: 256-276.

Knowles, J., Persico, N. & P. Todd. 2001. "Racial Bias in Motor Vehicle Searches." *Journal of Political Economy*, 109 (1): 203-229.

Landes, W. M. 1971. "An Economic Analysis of the Courts." *Journal of Law and Economics*. 14: 61-108.

Linder, D. O. 1996. "Juror Empathy and Race." *Tennessee Law Review*. 63: 887-916 (1996).

Miethe, T. D. 1987. "Charging and Plea Bargaining Practices under Determinate Sentencing." *Journal of Criminal Law & Criminology*, 78: 155-176.

Miller, M. L. 2004. "Domination & Dissatisfaction: Prosecutors as Sentencers." *Stanford Law Review,* 56: 1211-1276.

Oaxaca, R. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review*, 14: 693-709.

Oaxaca, R. & Ransom, M. 1994. "On Discrimination and the Decomposition of Wage Differentials." *Journal of Econometrics*, 61 (1): 5-21.

Oaxaca, R. & Ransom, M. 1999. "Identification in Detailed Wage Decompositions." *Review of Economics and Statistics*, 81 (1): 154-157.

Posner, R. A. & Yoon, A. H. 2011. "What Judges Think of the Quality of Legal Representation." *Stanford Law Review*, 63: 317-350.

Prescott, J., 2006. "Empirical Evidence of Prosecutorial Charging Manipulation," *mimeo.*

Rehavi, M. & Starr, S. 2012. "The Effects of Sentencing Law Reform on Racial Disparity in Charging," *mimeo.*

Schulhofer, S. J. & I. Nagel. 1997. "Plea Negotiations Under the Federal Sentencing Guidelines." *Northwestern University Law Review*, 91: 1284-1316.

Scott, R. W. 2012. "Inter-Judge Sentencing Disparity After Booker: A First Look." *Stanford Law Review*, 63 (1): (forthcoming).

Shermer, L. O. & B. D. Johnson. 2010. "Criminal Prosecutions: Examining Prosecutorial Discretion and Charge Reductions in U.S. Federal District Courts". *Justice Quarterly* 27: 394-395.

Skolnik, S. 1999. "Grand Jury: Power Shift?" *Legal Times*, April 12.

Smith, J. and P. Todd. 2005. "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of Econometrics,* 125: 305-353.

Starr, S. B. 2012. "Gender Disparity in Federal Criminal Cases," Working Paper.

Stith, K. 2008. "The Arc of the Pendulum: Judges, Prosecutors, and the Exercise of Discretion." *Yale Law Journal*, 117:1420-1497.

United States Sentencing Commission. 2010. "Demographic Differences in Federal Sentencing Practices" [hereinafter USSG updated report].

United States Sentencing Commission. 2011. "Report to Congress: Mandatory Minimum Penalties in the Federal Criminal Justice System", Oct. 2011.

Table 1: Summary Statistics

	Filing Sample [Cases Initiated FY 07-09]			Sentencing Sample [Cases Sentenced FY 07-09]		
	Obs	Mean	BlkMean	Obs	Mean	BlkMean
Demographics/SES						
Black	36,067	0.45	N/A	32,346	0.45	N/A
Educ Cat 1: Dropout	22,239	0.35	0.43	32,346	0.35	0.43
Educ Cat 2: HS Dipl.	22,239	0.19	0.17	32,346	0.19	0.18
Educ Cat 3: GED	22,239	0.21	0.21	32,346	0.20	0.21
Educ Cat 4: College	22,239	0.24	0.19	32,346	0.25	0.19
Age	36,067	36.2	32.4	30,830	35.8	32.2
County Per Cap. Inc.	36,067	21,053	21,242	32,346	21,091	21,259
County Poverty %	36,067	14.2	14.7	32,346	14.2	14.7
County Unempl.%	36,067	4.0	4.1	32,346	4.0	4.1
County Crime per 1000	36,067	6.2	7.4	32,346	6.3	7.5
Case Attributes						
Appointed Counsel	13,996	0.70	0.85	13,024	0.68	0.84
Multi-Defendant Case	36,067	0.26	0.25	32,346	0.26	0.25
Property/Fraud Offense	36,067	0.33	0.25	32,346	0.33	0.24
Regulatory/Other Offense	36,067	0.11	0.06	32,346	0.10	0.05
Weapons Offense	36,067	0.42	0.55	32,346	0.44	0.57
Violent Offense	36,067	0.13	0.14	32,346	0.12	0.14
Criminal History						
Category 1 [least]	22,624	0.35	0.21	32,346	0.35	0.20
Category 2	22,624	0.09	0.09	32,346	0.09	0.09
Category 3	22,624	0.15	0.17	32,346	0.14	0.17
Category 4	22,624	0.13	0.17	32,346	0.12	0.17
Category 5	22,624	0.09	0.12	32,346	0.09	0.12
Category 6 [most]	22,624	0.20	0.24	32,346	0.20	0.26
Charge Severity						
Statutory Max	35,107	191.9	210.4	31,316	192.6	210.7
[S.D.]		[141.6]	[156.0]		[141.5]	[155.6]
Guidelines	35,086	47.7	56.2	29,910	48.0	55.9
[S.D.]		[65.6]	[76.3]		[65.4]	[74.7]
Past Mean Sentence	26,605	48.08	57.7	24,267	48.4	57.6
[S.D.]		[42.3]	[47.0]		[41.9]	[45.7]
Mandatory Min Dummy	35,200	0.13	0.18	31,385	0.13	0.18
[S.D.]		[0.33]	[0.39]		[0.33]	[0.39]
Case Outcomes						
Non-Petty Conviction	25,181	0.93	0.93	32,346	1	1
[S.D.]		[0.25]	[0.25]		[0]	[0]
Incarceration Dummy	22,905	0.81	0.88	32,346	0.82	0.89
[S.D.]		[0.39]	[0.33]		[0.39]	[0.31]
Prison Sentence Length Incl. Zeros [S.D.]	22,905	53.7 [71.0]	69.4 [80.2]	32,346	58.2 [76.7]	74.6 [85.3]
Prison Sentence Length If Incarcerated	18,522	66.5 [73.4]	79.0 [81.1]	26,485	71.1 [79.2]	84.0 [86.1]

Table 2: Initial Charge Severity Measures

	Mandatory Min Dummy	Log Statutory Max	Log Guideline Sentence	Log Past Mean Sentence
Black	2.04**	0.092**	0.058**	0.087**
	[0.11]	[0.009]	[0.009]	[0.0010]
Age	0.981**	-0.003**	-0.002**	-0.003**
	[0.002]	[0.0004]	[0.0004]	[0.0005]
Multiple Defendant	2.33**	0.237**	0.114**	0.256**
	[0.16]	[0.013]	[0.016]	[0.016]
County Poverty	0.98	0.004	0.003	0.006
	[0.01]	[0.002]	[0.002]	[0.003]
County Unempl.	1.03	-0.010	-0.011	-0.012
	[0.04]	[0.008]	[0.007]	[0.008]
Log County Income	0.70	0.058	-0.029	-0.018
	[0.13]	[0.040]	[0.038]	[0.044]
County Crime Rate	1.01	0.000	-0.002	0.003
	[0.01]	[0.002]	[0.002]	[0.003]
N	32875	35403	35381	26812

Col. 1 gives odds ratios from a logistic regression; Cols. 2-4 contain OLS coefficients. All include district and arrest-offense fixed effects. Standard errors clustered by arrest offense-district. * p<0.05, **p<0.01

Table 3: Alternate Specifications for Charge Regressions: Black Coefficients

	Mandatory Min Dummy	Log Statutory Max	Log Guideline Sentence	Log Past Mean Sentence
1. Main Specification	2.04**	0.092**	0.058**	0.087**
	[0.11]	[0.009]	[0.009]	[0.010]
2a. FY 2007 Arrests [Main Specif.]	1.55**	0.064**	0.023	0.054**
	[0.18]	[0.018]	[0.018]	[0.018]
2b. FY 2007: Add Police Notes Flags	1.65**	0.073**	0.030	0.058**
	[0.19]	[0.018]	[0.018]	[0.018]
3. Exclude Prearrest Indictment Cases	2.08**	0.089**	0.061**	0.093**
	[0.13]	[0.012]	[0.011]	[0.012]
4. South Only	1.96**	0.091**	0.056**	0.087**
	[0.15]	[0.013]	[0.013]	[0.014]
5a. Sentencing Sample: Main Specif.	1.92**	0.087**	0.038**	0.084**
	[0.12]	[0.009]	[0.009]	[0.01]
5b. Sentencing Sample: Add Educ	1.92**	0.088**	0.037**	0.083**
	[0.12]	[0.009]	[0.009]	[0.01]
5c: Sentencing Sample: Add Crim Hist	1.87**	0.078**	0.034**	0.058**
	[0.11]	[0.009]	[0.009]	[0.09]
6a. Cases w/ Counsel Type [Main Specif.]	2.22**	0.122**	0.101**	0.108**
	[0.20]	[0.017]	[0.016]	[0.017]
6b. Add Counsel Type	2.11**	0.110**	0.848**	0.099**
	[0.18]	[0.017]	[0.016]	[0.017]
7. High Crime Counties Only	2.17**	0.085**	0.049**	0.061
	[0.19]	[0.013]	[0.021]	[0.013]

Odds ratios on black indicator (column 1) and OLS coefficients (columns 2-4) for variants of the Table 2 regressions, as described in the row labels. Standard errors clustered by arrest offense-district in brackets. * p<0.05, **p<0.01

Table 4: Selection into Prison Sample: Conviction and Incarceration

	Non-Petty Conviction	Incarceration
Black	0.957	1.097
	[0.054]	[0.055]
N	31,927	32,319

Odds ratios from logistic regressions including district, arrest offense, age, multi-defendant flag, and county income, poverty, unemployment, and crime rate; Col. 2 also includes criminal history and education. Standard errors clustered by arrest offense-district in brackets.
* p<0.05, **p<0.01

Table 5: Mean & RIF Decile Decompositions of Racial Disparity in Log Sentence Length

Panel A: No Controls for Initial Charging [N=25,695]

	Mean	10	30	50	70	90
Raw Gap	0.448**	0.424**	0.443**	0.470**	0.391**	0.440**
	[0.028]	[0.046]	[0.033]	[0.030]	[0.033]	[0.033]
Unexplained	0.096**	-0.052	0.050*	0.125**	0.068**	0.199**
	[0.016]	[0.039]	[0.021]	[0.017]	[0.020]	[0.027]
Explained:						
Arrest Offense	0.179**	0.249**	0.197**	0.174**	0.162**	0.112**
	[0.018]	[0.027]	[0.020]	[0.017]	[0.017]	[0.015]
Criminal History	0.156**	0.207**	0.177**	0.142**	0.128**	0.121**
	[0.011]	[0.016]	[0.013]	[0.011]	[0.012]	[0.012]
Education	-0.000	0.003	-0.002	-0.004	0.000	-0.001
	[0.002]	[0.004]	[0.003]	[0.002]	[0.003]	[0.003]
Other	0.017	0.017	0.021	0.033*	0.032	0.009
	[0.013]	[0.020]	[0.013]	[0.014]	[0.017]	[0.018]

Panel B: Mandatory Minimum Charge Dummy Added [N=25,695]

	Mean	10	30	50	70	90
Raw Gap	0.448**	0.424**	0.443**	0.470**	0.391**	0.440**
	[0.028]	[0.045]	[0.033]	[0.030]	[0.033]	[0.033]
Unexplained	0.048**	-0.085*	0.011	0.077**	0.009	0.146**
	[0.016]	[0.039]	[0.021]	[0.017]	[0.019]	[0.028]
Explained						
Mand Min	0.068**	0.047**	0.055**	0.067**	0.084**	0.075**
	[0.007]	[0.006]	[0.006]	[0.007]	[0.009]	[0.009]
Arrest Offense	0.168**	0.241**	0.187**	0.162**	0.148**	0.099**
	[0.017]	[0.027]	[0.019]	[0.016]	[0.016]	[0.014]
Criminal History	0.159**	0.209**	0.179**	0.145**	0.132**	0.124**
	[0.011]	[0.016]	[0.013]	[0.011]	[0.012]	[0.012]
Education	0.000	0.004	-0.001	-0.003	0.001	-0.001
	[0.002]	[0.004]	[0.003]	[0.002]	[0.002]	[0.003]
Other	0.006	0.010	0.012	0.022	0.018	-0.003
	[0.013]	[0.020]	[0.013]	[0.014]	[0.016]	[0.018]

Table 5, cont.

Panel C: Mandatory Minimum Dummy and Log Stat Max Added **[N=25,695]**

	Mean	10	30	50	70	90
Raw Gap	0.448**	0.424**	0.443**	0.470**	0.391**	0.440**
	[0.028]	[0.045]	[0.033]	[0.030]	[0.033]	[0.033]
Unexplained	0.048**	-0.086*	0.011	0.077**	0.008	0.145**
	[0.016]	[0.039]	[0.021]	[0.017]	[0.018]	[0.027]
Explained:						
Mand Min	0.027**	0.022*	0.017**	0.016**	0.020**	0.026**
	[0.005]	[0.009]	[0.005]	[0.004]	[0.005]	[0.006]
Log Stat Max	0.035**	0.022**	0.033**	0.044**	0.055**	0.042**
	[0.006]	[0.007]	[0.006]	[0.008]	[0.009]	[0.008]
Arrest Offense	0.169**	0.242**	0.189**	0.165**	0.151**	0.102**
	[0.017]	[0.027]	[0.019]	[0.016]	[0.015]	[0.013]
Criminal History	0.162**	0.211**	0.183**	0.150**	0.138**	0.129**
	[0.011]	[0.016]	[0.013]	[0.012]	[0.012]	[0.012]
Education	0.001	0.004	-0.001	-0.002	0.002	0.000
	[0.002]	[0.004]	[0.003]	[0.002]	[0.002]	[0.003]
Other	0.005	0.009	0.012	0.021	0.017	-0.004
	[0.013]	[0.020]	[0.013]	[0.014]	[0.016]	[0.017]

Oaxaca pooled decompositions of the black-white gap at the mean and in the RIF at the deciles. "Other" contains: district fixed effects, multiple defendant flag, age, and county unemployment, poverty, and crime rates as well as per capita income. The reported deciles correspond to sentences of 12, 27, 46, 77 and 180 months, respectively. Standard errors clustered by arrest-offense-district are in brackets. * p<0.05, **p<0.01

Table 6: Racial Disparity in Log Sentence Length Explained by use of Mandatory Minimums, Alternate Specifications

	Mean	10	30	50	70	90	Observations
Main Specification [Table 5b]	0.068** [0.007]	0.047** [0.006]	0.055** [0.006]	0.067** [0.007]	0.084** [0.009]	0.075** [0.009]	25695
Exclude Pre-Arrest Indictments	0.079** [0.008]	0.051** [0.006]	0.064** [0.007]	0.078** [0.008]	0.096** [0.010]	0.090** [0.011]	14696
South Only	0.082** [0.009]	0.056** [0.008]	0.069** [0.008]	0.088** [0.010]	0.101** [0.012]	0.076** [0.010]	13067
High Crime Counties	0.071** [0.008]	0.053** [0.008]	0.057** [0.007]	0.072** [0.009]	0.093** [0.011]	0.073** [0.010]	13493
Sub-sample: Counsel Type Recorded	0.102** [0.010]	0.067** [0.009]	0.080** [0.009]	0.098** [0.010]	0.130** [0.014]	0.122** [0.016]	9900
Includes Counsel Type Control	0.102** [0.010]	0.067** [0.009]	0.080** [0.009]	0.098** [0.010]	0.130** [0.014]	0.122** [0.016]	9900
Sub-sample: Police Notes Available	0.062** [0.008]	0.038** [0.006]	0.054** [0.007]	0.064** [0.008]	0.079** [0.011]	0.069** [0.010]	11265
Includes Police Notes	0.061** [0.008]	0.038** [0.006]	0.052** [0.007]	0.062** [0.008]	0.076** [0.011]	0.070** [0.010]	11265
Mandatory Minimum Conviction:							
Authors' Coding Scheme	0.062** [0.007]	0.038** [0.005]	0.047** [0.005]	0.059** [0.006]	0.077** [0.009]	0.073** [0.009]	25695
USSC Coding	0.140** [0.011]	0.070** [0.007]	0.094** [0.008]	0.130** [0.010]	0.194** [0.015]	0.242** [0.020]	25661
USSC Coding: Includes Drugs and Child Pornography Cases	0.113** [0.015]	0.085** [0.012]	0.128** [0.018]	0.131** [0.018]	0.116** [0.016]	0.101** [0.014]	58861

Each cell is from a separate Oaxaca pooled decompositions of the black-white gap at the mean and in the RIF at the deciles. Each cell contains the percentage points of the black-white sentence gap that can be explained by the presence of a charge carrying a mandatory minimum sentence. All decompositions include controls for arrest offense, criminal history, education, district fixed effects, multiple defendant flag, age, and county unemployment, poverty, and crime rates as well as per capita income. Standard errors clustered by arrest-offense-district are in brackets. * p<0.05, **p<0.01

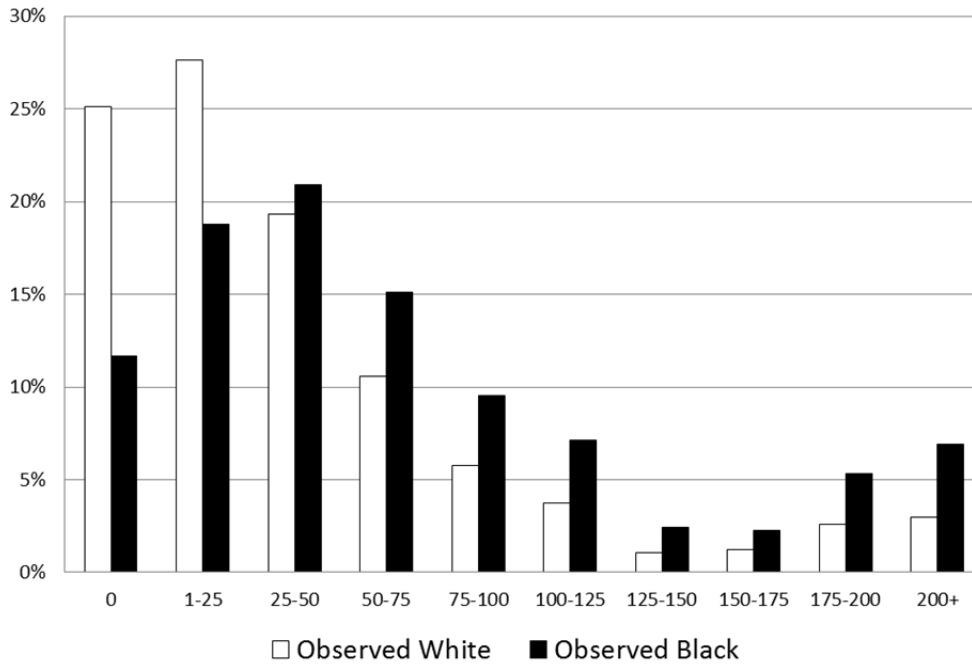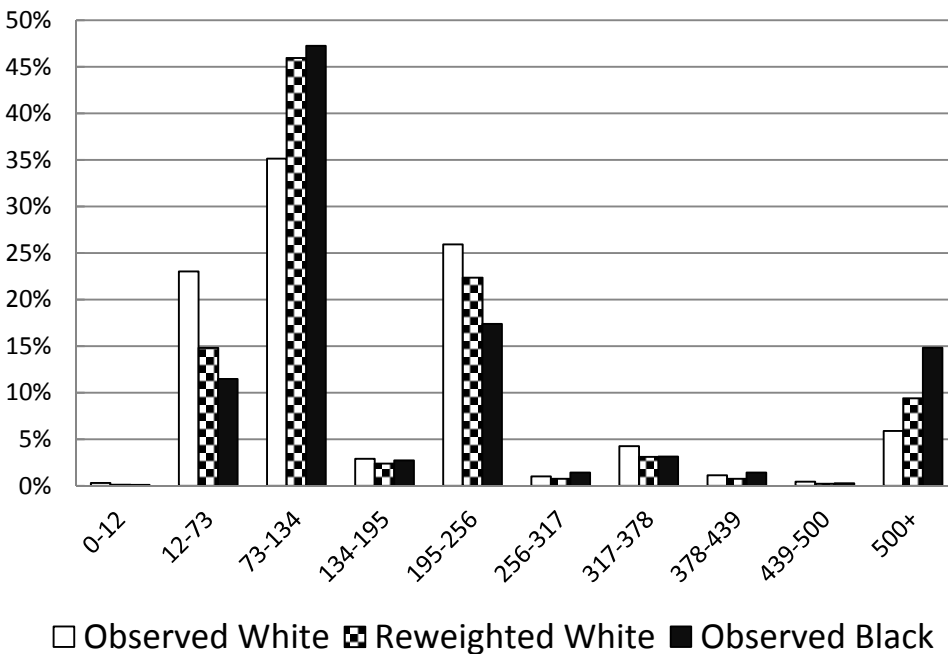**Figure 1: Distribution of Prison Sentences [Months][47]**

Observed White   Observed Black

Figure 2a: Distribution of Observed & Reweighted Charge Severity: Statutory Max[48]

Observed White   Reweighted White   Observed Black

---

[47]All ranges are up to and including the upper limit, and commence just above the lower limit (except for the first bin of each graph, which starts at 0). For instance, in Fig. 1, an observation with a 1-month sentence is placed in the 0-1 bin; 1 month plus one day is placed in the 1-25 bin.

[48] Reweighting of white distributions in Figs. 2a-2c is based on black endowments of arrest offense, district, age, multi-defendant flag, and county poverty, unemployment, income, and crime rate.

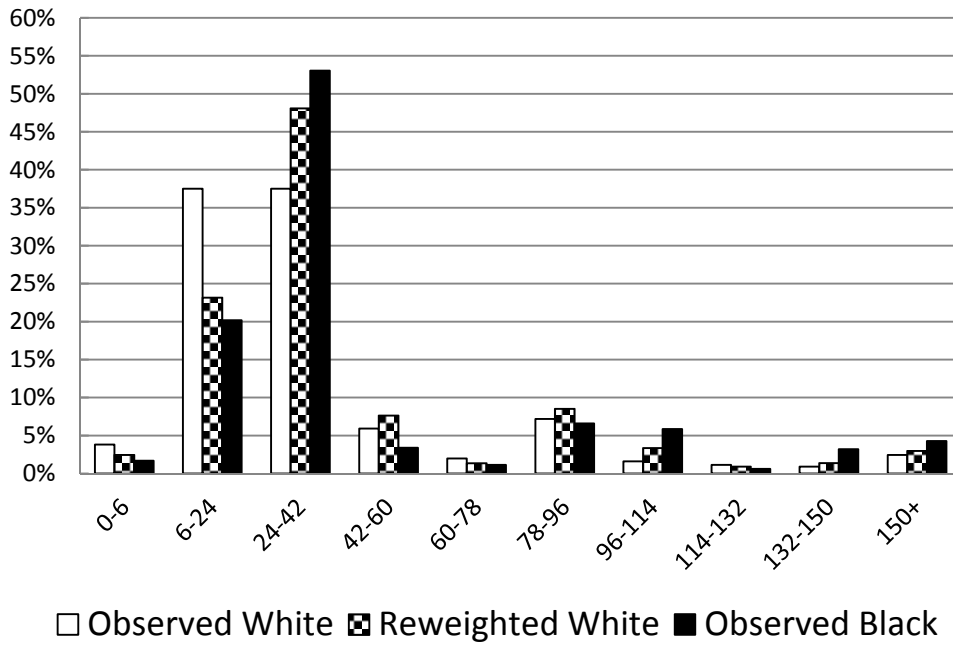**Figure 2b: Distribution of Observed & Reweighted Charge Severity: Guideline Scale**

□ Observed White ▣ Reweighted White ■ Observed Black

Figure 2c: Distribution of Observed & Reweighted Charge Severity: Past Mean Sentence
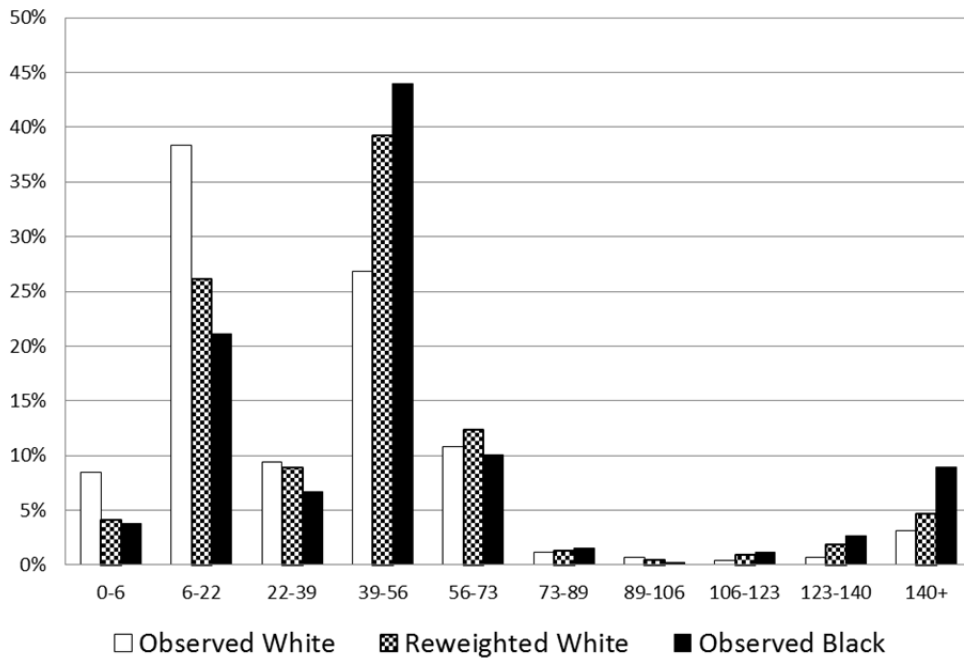

□ Observed White ▣ Reweighted White ■ Observed Black

Figure 3a: RIF Decomposition of Black-White Sentence Gaps at the Deciles of the Log Sentence Distribution: Pre-Charge Controls Only [Table 5a][49]

Black-White Gap   Total Explained (without Charging)
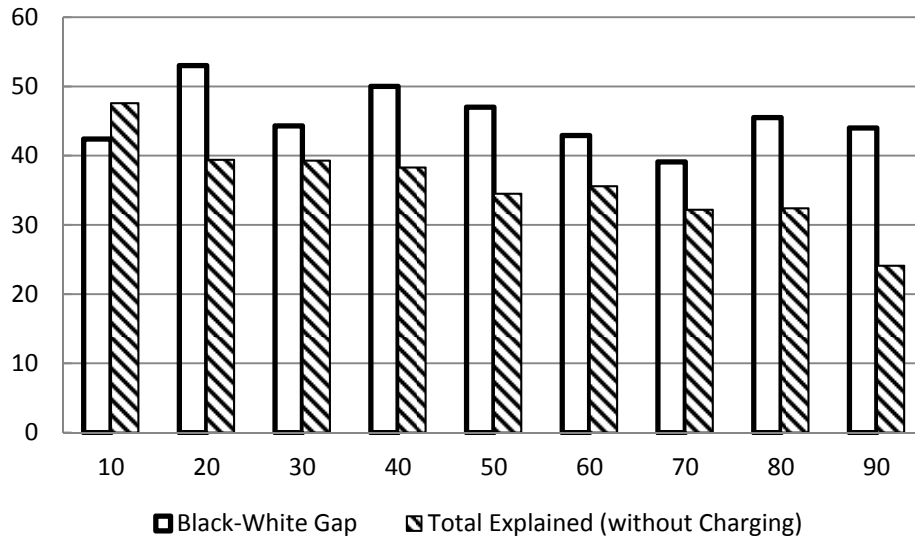
Figure 3b: RIF Decomposition of Sentencing Disparities at the Deciles of the Log Sentence Distribution: Mandatory Minimum Charge Indicator Added [Table 5b]

Black-White Gap   Total Explained   Explained: Mandatory Minimums

---
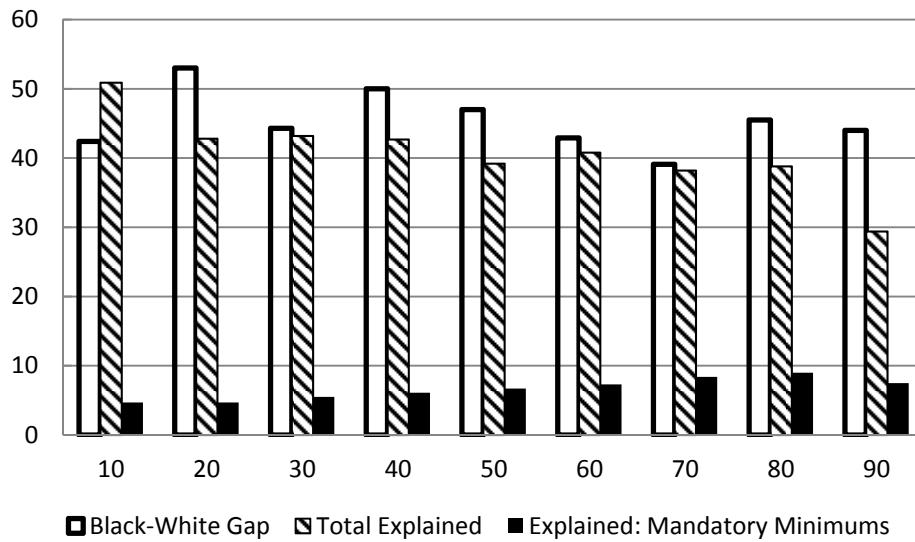
[49] Decomposition in Fig. 3a is based on arrest offense, district, criminal history, education, age, multi-defendant flag, and county poverty, unemployment, income, and crime rate. In Fig. 3b, it also includes the mandatory minimum charge indicator.

Figure 4: Distribution of Observed Prison [Months] & DFL Reweighting on Arrest, Mandatory Minimum, & Pre-Arrest Characteristics[50]

---
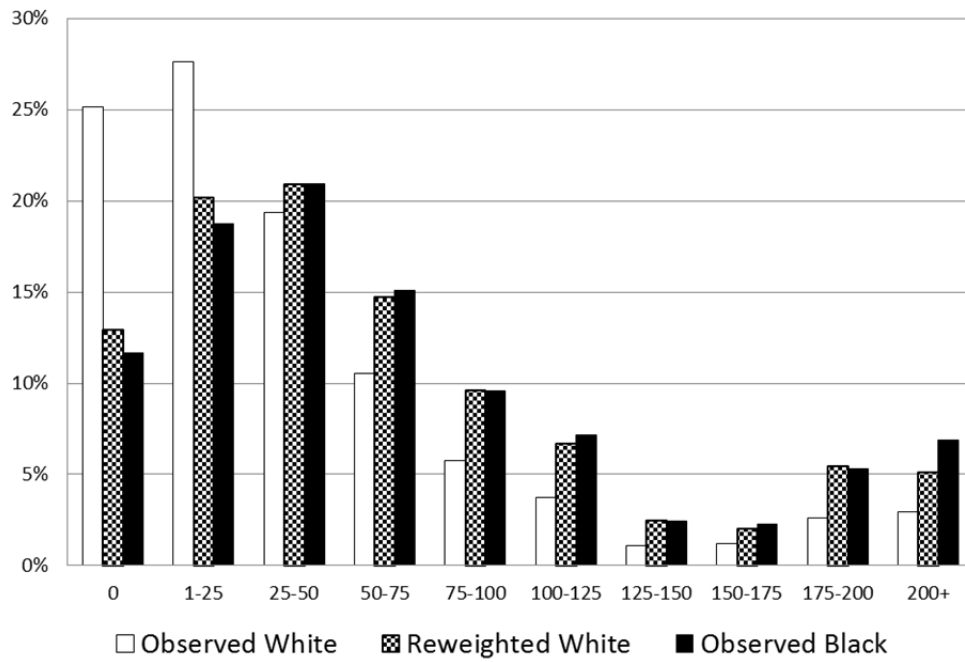
[50] Reweighting is based on mandatory minimum dummy, arrest offense, district, criminal history, education, age, multi-defendant flag, county poverty, unemployment, income, and crime rate.

DATA APPENDIX

Construction of the Linked Dataset:

The primary data for this analysis are restricted data collected by the Bureau of Justice Statistics (BJS) and available to researchers via the National Archive of Criminal Justice Data.[1] Restricted-use dyadic linking files from BJS link case records across agencies as they go through the criminal justice process, enabling cases to be followed from arrest through to sentencing. Files are linked across the following agencies: the U.S. Marshals' Service (USMS), the Executive Office for the U.S. Attorneys (EOUSA), the Administrative Office of the U.S. Courts (AOUSC) and the U.S. Sentencing Commission (USSC).

USMS collects arrest-stage data from every federal law enforcement agency as well as local officials that transfer arrestees to federal custody. These data are the source of the following fields used in the main specifications and/or robustness checks or to define the sample: arrest offense, race, age, gender, police notes describing the offense, U.S. citizenship status, arrest date, the federal judicial district, and the arresting agency. EOUSA collects a variety of charging and investigation-related information recorded by prosecutors. We only used EOUSA's data for limited purposes: to identify the date charges were filed, for the initial analysis of whether a suspect file opened by a prosecutor resulted in charges being filed in district court (instead of being resolved by magistrate), to determine whether the seizure of drugs upon arrest was mentioned in the prosecutor's investigation-stage records, and as part of the linking pathways described below to connect the arrest data to later stages in the process.

The main charge severity analysis, which used only cases filed in district court, relied on the more detailed charge information collected by the AOUSC from district courts. The AOUSC data identify five initial charges in each case by the U.S. Code subsection corresponding to the crime. The data also list up to five final charges and the disposition of each charge, and this information was used to identify the charges of conviction. AOUSC data were also used to identify the number of defendants in each case and the defense counsel type.

Sentencing information was taken from the USSC data. USSC is also the only source of information on the defendant's criminal history, education level, and Hispanic ethnicity, and its data were also used in robustness checks as an alternate identifier of the mandatory minimum sentence applicable to the crime of conviction.

The linking algorithm is dyadic, such that agencies' files must be linked in the following order: from USMS to EOUSA investigation and case files, from EOUSA to AOUSC case files, and from AOUSC to USSC.[2] The main estimation sample for the charge severity analysis is limited to cases that could be linked from the USMS records through the EOUSA files to the AOUSC files. Approximately 81% of cases could be linked from the USMS arrest files to EOUSA investigation files.[3] Among cases that successfully linked and resulted in charges being filed in district court,[4] 92% could be linked to AOUSC initial charging data and were thus used in the main charging analysis. Conditional on arrest offense, district, and age (the covariates from the main analysis that were observable in the USMS data), there were no significant racial differences in either of these link-through rates.[5]

Cases used for the sentence length analysis had to be further linked through to the USSC data, which covers only cases sentenced for non-petty offenses (defined as offenses carrying a

---

[2] More specifically, there are multiple types of "Standard Analysis Files" (SAFs) from EOUSA ("Matters Out" files on criminal suspects, "Cases In" files on cases filed in district court in each year, and "Cases Out" files on cases terminated in district court in each year) and from AOUSC ("Cases In" and "Cases Out," for cases filed and terminated each year, respectively). BJS's linking algorithm offers two possible linking pathways for cases handled by district courts, one that connects the "Cases In" files between EOUSA and AOUSC, and one that connects the "Cases Out" files; each pathway is then supplemented with a separate set of intra-agency links so that all SAFs from each agency could be used whenever possible. Both pathways were used to maximize total linking rates. SAFs were appended across years before being linked across agencies. The Cases In and Cases Out files for each agency contain redundant fields; Cases Out is simply an updated version created when the case is terminated that adds the terminal-stage information (e.g., the disposition of each charge). In the charging-stage analysis, we used the AOUSC Cases In file as the preferred data source (except in unusual cases where only Cases Out was available due to an intra-agency linking problem), because that was the source that was available for cases not yet terminated by the end of FY 2009. The analysis of charges of conviction was based on AOUSC Cases Out.

[3] All linking percentages are within a pool of cases that tracks the main sample used for analysis as closely as possible without using fields that are only available for cases that successfully link. Thus, this percentage is within cases identified by USMS as black or white U.S. citizen males, excluding drug and immigration crimes and cases from districts lacking in minimal racial diversity; these exclusion criteria are detailed below. Cases that did not link at this stage could have been due to failures of the linking algorithm, due to cases being handled by some prosecutorial unit other than a U.S. Attorney's Office, or due to the prosecutor choosing not to open an investigation file. According to EOUSA, the U.S. Attorney's offices handle approximately 95% of federal prosecutions; the others are handled by specialized units within the Department of Justice.

[4] As explained below, our main analysis focused on cases filed in district courts rather than being resolved by magistrates.

[5] This is based on logistic regressions of an indicator of linking failure within the same pool of cases for which linking percentages are reported, with standard errors clustered at the offense-district level, the same level used in the main analysis.

statutory *maximum* penalty of more than six months' imprisonment, although the actual sentence could be less). We first used the AOUSC conviction data to identify which filed cases resulted in non-petty convictions (and analyzed racial disparity at that stage; see Table 4).   Among those non-petty conviction cases, 96% linked through to the Sentencing Commission data.[6] Conditional on all of the covariates used in the main filing-stage analysis (arrest offense, district, age, multi-defendant flag, and county poverty, unemployment, per capita income, and crime rate), there were, again, no significant differences by race in link-through rates. Moreover, there is no theoretical reason to expect the imperfections in any of the linking algorithms to bias the estimates of black-white racial disparity in charging or sentencing.

Sample Restrictions:

Timing: The charging analysis sample was limited to individuals who were initially charged or arrested (whichever was later) between fiscal years 2007 and 2009, inclusive (October 1, 2006 through September 30, 2009).[7] These are the most recent years of data available from BJS. To be included in the sentencing analysis an individual had to be sentenced between fiscal years 2007 and 2009, inclusive. This approach avoided (as opposed to requiring all dates to be in the sample period) avoided disproportionately excluding cases that were processed more slowly, which might have differed from faster-processed cases in terms of case type, complexity, or choices made by the parties.   The result of the approach is overlapping but distinct analysis samples at each stage. Approximately 61% of the filing sample is also found in the sentencing sample,[8] and approximately 68% of the (smaller) sentencing sample is found in the filing sample.[9] Certain filing-stage robustness checks used fields drawn from sentencing data and thus were conducted within the sentencing sample, i.e., with cases filed on average somewhat earlier; likewise, the robustness checks using the police notes string field were

---

[6] This linking rate is based on convictions from FY 2007-08; convictions from FY 2009 had a lower linking rate because not all of them were sentenced by the end of FY 2009. As explained below, the sentencing-stage sample was defined in terms of sentencing date to avoid truncation problems.

[7] In cases in which no formal indictment was brought (meaning the parties agreed to allow the case to proceed on the "complaint" underlying the arrest warrant), the arrest date was treated as the relevant date.

[8] Approximately 4% of the filing-sample cases were lost to linking failures before the sentencing stage (as explained above), approximately 3% did link but had missing data on either criminal history or education (the controls added at the sentencing stage), and approximately 7% are cases in which there was no conviction of a non-petty offense. This leaves approximately 25% of the filing-sample cases that were presumably sentenced after the end of FY 2009.

[9] The balance of the sentencing sample consists of cases filed before FY 2007. Ninety-seven percent of the cases in the sentencing sample were, however, filed after the Supreme Court's January 2005 decision in *United States v. Booker.*

conducted within a subset of the earliest cases, because the field is only found in the USMS data for arrests from 2007 or before. These checks drawing on earlier data show slightly smaller disparities even before additional variables are added to the specification, suggesting that charging disparity has grown somewhat over time (although change over time is not a focus of this study).

Defendant Type: Because most federal felonies render non-citizens deportable (often automatically), the effective severity of charges in non-citizens' cases is not readily comparable to that of citizens, and they were accordingly excluded from the main analysis samples (although added back in robustness checks). U.S. citizenship was identified based on the USMS data. Women were also excluded based on USMS gender data. People of races other than black and white (as defined below) were also excluded.

Case Type: Due to the deportation stakes, as well as the very different ("fast-track") procedural framework many districts apply to immigration cases, we also eliminated immigration cases (almost all of which would have been excluded in any event due to the citizenship requirement).

Drug cases were also excluded due to the lack of reliable drug quantity information at the charging and arrest stages; quantity is crucial to understanding the severity of drug charges. EOUSA's data do include a "quantity seized" field representing the prosecutor's recording of the amount of drugs seized at arrest; this field is initially recorded in the suspect investigation file, prior to charging, and thus might serve as arrest-stage data (although its recording may involve some prosecutorial discretion). However, we discovered drastic changes in the apparent quantity distribution in this field from 2003 to 2004 (the year EOUSA adopted a new data entry system) as well as large inconsistencies in quantity between this field and the sentencing-stage quantities recorded by USSC beginning in 2004, suggesting that the problem is with the post-2003 EOUSA data.[10] We could identify no reason for the shifts other than data entry problems, and the problems were not uniformly applicable nor confined to particular drug types or districts; there was no way to identify which specific cases had the wrong quantities recorded.

Child pornography cases were excluded due to our inability to distinguish between certain key sub-provision of the charging statutes listed by AOUSC. In particular, simple

---

[10] Although USSC records drug quantity, this field is the product of sentencing fact-finding and influenced by charging and plea negotiations, and thus could not be treated as an exogenous measure of the quantity seized at arrest.

possession cases could not be distinguished from receipt or distribution cases, a distinction critical to the application of a mandatory minimum sentence.

Both drug cases and child pornography cases could be included in the broader sample used in the Table 6, Row 11 decomposition, because the mandatory minimum indicator used was based on the USSC data for the crime of conviction; the USSC data include the actual mandatory minimum recorded by the judge, rather than just the statute of conviction, so there was no need to resolve legal ambiguities in the AOUSC coding or to know the quantity of drugs alleged by the prosecutor.

All case type exclusions were based exclusively on the USMS arrest code, which is based on the arresting officer's characterization of the *principal* offense in the case--under federal sentencing law, the principal (most severe) offense is the main and usually the only driver of the ultimate sentence. Defining the sample based on the arrest stage data alone (rather than the nature of subsequent charges in the case) avoided potentially serious sample selection issues that could have emerged had the exclusions been based on the prosecutor's discretionary decisions.[11] Because criminal cases sometimes involve multiple types of criminality, this approach did not eliminate the possibility of secondary offense conduct related to the excluded categories. In the case of immigration and child pornography, this was rare: 0.02% and 0.2% of the cases in the sample involved any immigration and child pornography charges, respectively. However, a drug charge was brought in 8.8% of the cases, eighty percent of which were principally coded as weapons cases; a drug nexus is often the basis for federal criminal jurisdiction over gun possession cases. Robustness checks addressing the potential role of drugs in the cases are detailed in the paper.

Cases with arrest codes indicating a reason for detention other than a criminal offense (material witness warrants and violations of the conditions of parole or probation) were also excluded from the sample.

Other Exclusions: The sample was also limited to the 50 US States and the District of Columbia; the territorial districts (Puerto Rico, Guam, U.S. Virgin Islands) were excluded. All the analyses use district fixed effects, and thus the following districts with insufficient numbers

---

[11] For instance, if prosecutors disproportionately declined to pursue a gun charge in guns-and-drugs cases involving white defendants, white defendants in such cases would disproportionately disappear from the sample if it were defined based on the charges rather the arrest offense.

of black defendants were excluded: Hawaii, Idaho, Maine, Montana, North Dakota, South Dakota, Vermont, and Wyoming. The district exclusions reduced the sample size by 0.48%.

Construction of Key Independent Variables:

Race: Race is drawn from the USMS data, and is coded as white, black, Asian, Native, and Other/Unknown. The last three groups together constituted about 4% of the cases otherwise satisfying the sample requirements, and were dropped from the sample. The USMS does not include a separate category for Hispanic; rather, Hispanics are included within other racial groups. The USSC does record Hispanic ethnicity, but is only available for those sentenced for a guideline offense, therefore this field was used only in robustness checks.

Arrest Offense: There are 430 unique arrest offenses listed in the USMS data. However, within the main sample used for the filing analysis, over 95% of the cases fall under just 93 arrest offenses. The original arrest offense codes included many very similar offense descriptions, including some that were slightly more detailed versions of others (for instance, "vehicle theft" and "vehicle theft by bailee"). Often the more detailed ones were rarely used. Therefore, the smallest categories were combined with others that could describe the same legal offense, leading to 107 offense groups that were represented in the filing sample. No single numerical cutoff was used to determine when cases would be combined, because the combination depended on the legal assessment that the crimes were sufficiently similar. The results are robust to the use of the original offense codes.

Criminal History: Criminal history data are only available in the USSC data and are accordingly only available for those sentenced for guideline offenses. The variable used was the defendant's criminal history category, which ranges from 1 to 6 and forms the basis of the Guidelines sentencing grid. In 0.2% of the sentencing sample, this field was originally missing but could be calculated based on another Sentencing Commission field called "criminal history points," according to the rules laid out in the Guidelines.

Charge Severity Measures

The raw charge data consist of the statutory provision associated with each charge. We used this information to identify the statutory maximum and minimum sentence, the Guidelines recommended sentence, and the mean past sentence associated with each *combination* of charges

found. This required comprehensive legal research on the statutes and guidelines covering all federal offenses charged during the study period.

The core coding challenge was that the AOUSC charge fields are not always very specific—they might, for instance, refer to a particular statutory provision that contains two distinct subparagraphs with different sentencing schemes. We researched the most common ways in which these statutes are charged in order to be able to make realistic assumptions in the face of such ambiguities. In general, for instance, we assumed the defendant had no prior convictions of the exact same crime, thus avoiding special penalties that a few statutes apply to recidivists and focusing on the severity of the particular offense in question.[12] If none of the arrest offenses was homicide-related and no listed charge fell under a homicide statute, we assumed the defendant did not kill anyone (an aggravating factor in a large number of statutes in which death is a rare result, from violations of maritime rules to health care fraud). Similarly, we assumed that defendants in non-assaultive property or regulatory offense cases did not physically injure anyone.

When possible, we resolved ambiguities by reference to the other charges in the case, when the legal elements of those charges revealed additional facts that the prosecutor must have been alleging. In these cases we made exceptions to our default assumptions. For instance, suppose Charge 1 is a burglary offense that usually has a maximum sentence of 10 years, but has a 20-year maximum if someone is seriously injured in the course of the burglary. Charge 2 is an aggravated assault charge, with a 15-year maximum, in which aggravated assault is defined to require that serious injury be proven. In that case we would flag Charge 2 with a "serious injury" indicator, and that flag would trigger an enhancement to the coded statutory maximum for Charge 1, raising it to 20 years.

Implementing this approach required constructing a number of flags for every federal criminal statute. We constructed flags to indicate whether each of the following key facts were built into the required elements of the crime: death, injury, serious injury, drug crime, sex crime, fraud, official victim, minor victim, terrorist motive, an assault, use of a weapon, use of a gun specifically, a "crime of violence," obstruction of justice, taking a person for ransom, and

---

[12] This assumption is not the same as assuming that defendants have no prior criminal history at all; criminal history was a directly observable variable for sentenced cases and was used in the main sentencing specifications. However, prior convictions of the exact same offense can safely be assumed to be exceptional even among defendants with criminal history; indeed, most prior offenses on the records of federal defendants are not federal crimes at all, but state crimes.

whether the crime was a predicate offense for the crime of felony murder. For each statute, we also indicated any adjustments to the statutory or guidelines sentences that would be triggered by the presence of particular facts as identified by the flags for the other charges in the case. We followed this basic approach for each of the legal measures. Remaining ambiguities were resolved according to default assumptions that varied between the measures—these assumptions were generally designed to be conservative and err on the higher side for the statutory maximum and on the lower side for the statutory minimum and the guidelines measure. In the case of the statutory minimum, in practice this meant assuming *no* statutory minimum existed when faced with ambiguities; many criminal statutes ordinarily have no minimum except when some special circumstance is triggered.[13]

We chose to construct our own measure of the statutory maximum rather than use the existing AOUSC "severity" field, which is ostensibly based on the statutory maximum. The AOUSC coding appears to automatically be based on the very highest maximum contained *anywhere* in the statute cited, even when that maximum is only triggered by an exceptional circumstance that rarely applies. For instance, charges under 18 U.S.C. § 1347 (health care fraud) are coded by AOUSC as having a statutory maximum of life, even though that maximum only applies when the fraud leads to a death; the standard statutory maximum is ten years. Our approach uses, in our view, considerably more realistic assumptions.

Constructing a measure of the Guidelines sentence involved additional challenges. First, it required identifying the applicable Sentencing Guidelines for each case. These are not recorded by AOUSC. To that end, we conducted legal research analyzing the interrelationship of the Guidelines with every federal criminal statute. Second, unlike the statutory range, the Guidelines range is not solely determined by the charges; rather, it is heavily driven by sentencing fact-finding. However, because we are focused on the severity of charging, not subsequent sentencing fact-finding, we preferred a metric that, while Guidelines-based, was premised only on what the prosecutor actually charged: the elements of the crime. We accordingly based our Guidelines metric only on the offense level identified by the Guidelines assuming the elements of *all* charges brought were proven, but *no* additional findings of fact were made at sentencing. Since most sentencing facts are aggravating factors, this approach

---

[13] A detailed spreadsheet showing these flags and the assumed base statutory sentencing range and Guidelines offense level for each federal crime is available on request.

amounted to erring on the lower side, which was also our approach in the case of ambiguities in the AOUSC code that we could not resolve via legal research or information from other charges. Thus, the Guidelines metric is designed to err in the opposite direction from the statutory maximum measure.

The Guidelines define the "offense level"—a severity scale running from 1 to 43—associated with each offense. In order for the units of this measure to be comparable to the other metrics, this offense level had to be converted into an implied sentence length in months. Under the Guidelines, offense levels translate mechanically into sentence ranges based on a grid, with criminal history as the other axis. Our charging metric is by design blind to the defendant's actual criminal history—it reflects charge severity alone. Therefore, we used the same grid column for this translation in every case. We used the column corresponding to the highest criminal history category because it best preserves the distinctions between offense levels at the lower end of the scale and reduces the number of predicted zeros. Within the applicable sentence-range cell, we used the low end of the range.

The mean past sentence for each charge was constructed as follows: First, using AOUSC and Sentencing Commission data for fiscal years 2005 and 2006 (a period almost completely after *Booker* and immediately preceding our sample period), the average sentence corresponding to each statutory charge for white U.S. citizen defendants was calculated. The average was limited to cases sentenced only on one charge, so as to ensure that the sentence in fact corresponded to the charge. The past mean sentence was coded as missing for charges sentenced fewer than 30 times among this group in the reference period. The resulting averages were used to code the charges in the individual cases in the main 2007-09 sample. If a case had any charge that was not coded with a past mean sentence, the overall measure was missing and it was excluded from the analysis of this charging measure.

Once the severity of the individual charges were coded, we then combined them into the overall statutory minimum and maximum, guidelines measure, and past mean sentence measure for the case as a whole. Concurrent sentencing is the default federal rule for sentencing on multiple charges. This means that the total sentence is driven (often completely) by the most serious single charge. Therefore, unless the statute specifically required a consecutive sentence we assumed concurrent sentencing, such that the combined severity of the charges was the sentence associated with the charge carrying the highest sentence. When the statute specifically

requires consecutive sentencing, the sentences for the charges carrying consecutive sentences were added to the one for the most serious concurrent charge. This approach could underestimate the true way a judge would combine sentences, because secondary charges can sometimes increase the sentence even when consecutive sentences are not required by statute, particularly if the charges are based on distinct incidents or patterns of conduct. However, such cases cannot be identified from the available data, so the default assumption is the best approximation. We did, as described above, use information drawn from secondary charges to adjust the requisite statutory and guidelines sentences for the primary charge. Our approach to combining charges follows the method specified in the Sentencing Guidelines (see U.S.S.G. § 5G1.2).

We made two final adjustments to the combined charge severity measures. First, we imposed the statutory minimum and the sum of the individual-charge maximums as lower and upper bounds, respectively, on the Guidelines sentence, which also tracks sentencing law (see U.S.S.G. § 5G1.2). Second, we replaced zeros on the statutory maximum, guidelines, and mean sentence scales with half a month—half of the lowest nonzero values otherwise calculated—to reflect the fact that no criminal charge truly has zero severity, even if no incarceration is imposed. This adjustment affected only 0.05% of cases for the statutory maximum measure, 0.2% of cases for the guidelines measure, and 0.5% for the mean sentence measure.

After following the coding methods above, the statutory *minimum* for the combined charges in 87% of the cases in our sample was zero. We constructed a binary variable for whether any nonzero statutory minimum was given, and this was the basis for our main analyses of the role of mandatory minimums. This avoided the need to resolve certain ambiguities in the AOUSC charge coding, because some statutes provide differing lengths of mandatory minimum depending on the facts of the case.[14]

Conviction and Sentence Outcomes

We constructed dummy variables indicating whether the defendant was convicted of some crime, based on the final charge and disposition fields found in the AOUSC records. We also constructed a dummy for whether the defendant was convicted of a non-petty offense, also

---

[14] For the purpose of robustness checks, alternate versions of all of the charging measures were calculated that excluded any components of the total charge severity coming from drug crimes (affecting the 8.8% of the sample that involved both drug and other charges). To calculate these, the severity measures associated with the drug charges were set to zero before the combined charge severity was calculated.

based on AOUSC records; the classification of offenses is based on the statutory maximum, so we relied on our legal coding described above. Conviction of a non-petty offense is a prerequisite for inclusion in the Sentencing Commission data. We coded the severity of the charges of *conviction* using the same four measures described above (statutory maximum and minimum, guidelines severity, past mean sentence), by combining the information for all terminal charges for which the disposition field indicated a conviction. We coded as "true zeros" the charges that were dismissed or resulted in acquittal; the scale for non-dismissed charges began at half a month, just as it did at the initial charging stage.

Sentence Length: Sentence data were drawn from the Sentencing Commission and are therefore only available for those convicted of offenses covered by the sentencing guidelines. We truncated sentence lengths at 540 months, and replaced life sentences with that value. This length is longer than the highest non-life statutory maximum found in federal law (480 months), and corresponds approximately to the remaining life expectancy of an American of the sample-average age of 36 years. We coded probation- and fine-only sentences as a zero incarceration sentence, and they are included as zeros in Figures 1 and 4, which show the full sentencing distributions for blacks and whites. However, the main sentencing analysis focused on the log of the sentence length (in days) among those receiving at least some incarceration; the zeros were excluded from this analysis, and the threshold binary question whether some incarceration was given was assessed separately. The minimum observed non-zero sentence length was one day.

Data Sources:

U.S. Census Bureau, 2000 Census of Population and Housing, Summary File 3. Available at: http://www.census.gov/census2000/sumfile3.html

United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. Federal Justice Statistics Program: Paired-Agency Linked Files, 2009. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2011-11-11. ICPSR Study Number 30701-v3.

United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. Federal Justice Statistics Program: Arrests and Bookings for Federal Offenses, 2001-2009.[15] Ann

---

[15] Note that although the samples are limited to FY 2007-09 data, the sentencing sample consists of cases *sentenced* in 2007-09 and thus includes a few cases in which the arrest and/or charge may have been substantially earlier. We accordingly created a linked dataset going back to 2001 initially, and then

Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. ICPSR Study Numbers 24126-v2 (2011-03-08), 24145-v2 (2011-03-08), 24164.v2 (2011-03-08), 24181.v2 (2011-03-08), 24216.v2 (2011-03-08), 24199.v2 (2011-03-08), 24211.v2 (2011-03-08), 24226.v2 (2011-03-08), 24231.v2 (2011-03-08), 29428.v2 (2011-03-08), 30794-v1 (2011-07-22). Original Data Source: U.S. Marshals' Service ("USMS").

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Suspects in Federal Criminal Matters Concluded, 2001-2009 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. ICPSR Study Numbers: 24120-v2 (2011-03-08), 24139.v2 (2011-03-08), 24158.v2 (2011-03-08), 24175.v2 (2011-03-08), 24193.v2 (2011-03-08), 24210.v2 (2011-03-08), 24225.v2 (2011-03-08), 29424.v2 (2011-03-08), 30790.v1 (2011-06-03). Original Data Source: Executive Office of U.S. Attorneys ("EOUSA Matters Out").

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Defendants Charged in Criminal Cases Filed in District Court, 2001-2009 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. ICPSR Study Numbers: 24121-v2 (2011-03-08), 24140.v2 (2011-03-08), 24159.v2 (2011-03-08), 24176.v2 (2011-03-08), 24194.v2 (2011-03-08), 24211.v2 (2011-03-08), 24226.v2 (2011-03-08), 29426.v2 (2011-03-08), 30791.v1 (2011-06-03). Original Data Source: Executive Office of U.S. Attorneys ("EOUSA Cases In").

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Defendants in Federal Criminal Cases -- Terminated, 2001-2009 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. ICPSR Study Numbers: 24122-v2 (2011-03-08), 24141.v2 (2011-03-08), 24160.v2 (2011-03-08), 24177.v2 (2011-03-08), 24195.v2 (2011-03-08), 24212.v2 (2011-03-08), 24227.v2 (2011-03-08), 29433.v2 (2011-03-08), 30792.v1 (2011-06-03). Original Data Source: Executive Office of U.S. Attorneys ("EOUSA Cases Out").

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Defendants in Federal Criminal Cases Filed in District Court, 2001-2009 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. ICPSR Study Numbers: 24114-v2 (2011-03-08), 24133.v2 (2011-03-08), 24152.v2 (2011-03-08), 24169.v2 (2011-03-08), 24186.v2 (2011-03-08), 24204.v2 (2011-03-08), 24221.v2 (2011-03-08), 29402.v2 (2011-03-08), 30781.v1(2011-06-03). Original Data Source: Administrative Office of the U.S. Courts ("AOUSC Cases In").

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Defendants in Federal Criminal Cases in District Court -- Terminated, 2001-2009 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. ICPSR Study Numbers: 24115-v2 (2011-03-08), 24134.v2 (2011-03-08), 24153.v2 (2011-03-08), 24170.v2   (2011-03-08), 24187.v2   (2011-03-08), 24205.v2   (2011-03-08),

---

limited the samples within the larger dataset as described above. In addition, for the charging analysis, we defined the "past mean sentence" charge-severity metric on the basis of cases sentenced in 2005-06.

24222.v2 (2011-03-08), 29242.v2 (2011-03-08), 30784.v1 (2011-06-03). Original Data Source: Administrative Office of the U.S. Courts ("AOUSC Cases Out").

United States Department of Justice. Bureau of Justice Statistics. Federal Justice Statistics Program: Defendants Sentenced Under the Sentencing Reform Act, 2001-2009 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. ICPSR Study Numbers: 24127-v2 (2011-03-08), 24146.v2 (2011-03-08), 24165.v2 (2011-03-08), 24182.v3 (2011-03-08), 24200.v3 (2011-03-08), 24217.v3 (2011-03-08), 24232.v2 (2011-03-08), 29381.v2 (2011-03-08), 30795.v1 (2011-06-06). Original Data Source: U.S. Sentencing Commission ("USSC").

United States Department of Justice. Office of Justice Programs. Federal Bureau of Investigation. Uniform Crime Reporting Program Data: County-Level Detailed Arrest and Offense Data, 2007-2009. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. ICPSR Study Numbers: 30763-v1 (2012-01-25), 27644-v1 (2011-04-21), 25114-v1 (2009-07-31).

Appendix: Additional Tables and Figures

Table A1: Sample Selection Analyses and OLS Sentence Length Results

	[1] Filing in District Court [Odds Ratio]	[2] Non-Petty Conviction [Odds Ratio]	[3] Incarceration [Odds Ratio]	[4] Log Prison Sentence [OLS]
Black	1.062 [0.064]	0.957 [0.054]	1.097 [0.055]	0.097** [0.015]
Age	1.003 [0.002]	0.983** [0.002]	0.994** [0.002]	0.005** [0.001]
Multiple Defendant Case		0.848* [0.054]	1.259** [0.062]	0.223** [0.017]
Education Cat 2: High School Diploma			0.757** [0.038]	-0.061** [0.017]
Education Cat. 3: GED			0.917 [0.056]	0.019 [0.015]
Education Cat. 4: College			0.938 [0.047]	0.022 [0.018]
Criminal History Category II			2.017** [0.121]	0.137** [0.026]
Criminal History Category III			3.484** [0.210]	0.243** [0.022]
Criminal History Category IV			9.036** [0.797]	0.520** [0.024]
Criminal History Category V			14.565** [1.862]	0.705** [0.025]
Criminal History Category VI			25.421** [2.624]	1.036** [0.025]
Observations	47,680	31,927	32,319	26,484

Columns 1-3 report the odds ratios from logistic regressions for the probability of, respectively, (1) arrestees facing charges in district court, (2) district court defendants being convicted of a non-petty offense; and (3) those convicted of non-petty offenses being sentenced to incarceration. All regressions also contain arrest offense and district fixed effects; the Columns 2-4 regressions also include county poverty rate, unemployment rate, log per capita income, and crime rate. The base category for education is high school dropout. Columns 2 and 3 report the full set of coefficients for the regressions included in Table 4 of the paper. Column 4 provides OLS coefficients for a regression of the log of prison sentence length among defendants sentenced to incarceration. The control variables vary across regressions due to data availability. Omitted controls are not available for cases that are not filed or not sentenced. Standard errors clustered at the offense-district in brackets. * p<0.05, **p<0.01

Table A2: Black Coefficients from Charge Severity Regressions:
Additional Robustness Checks for Estimates in Table 2

	[1] Mand. Min Dummy	[2] Log Stat Max	[3] Log Guideline Sentence	[4] Log Past Mean Sentence
Main Specification [Table 2]	2.04**	0.092**	0.058**	0.087**
	[0.11]	[0.009]	[0.009]	[0.001]
Original Arrest Coding	2.037**	0.093**	0.059**	0.091**
	[0.120]	[0.009]	[0.009]	[0.010]
Sub-Sample Hispanic Recorded	1.926**	0.084**	0.037**	0.0878**
	[0.123]	[0.010]	[0.009]	[0.010]
Hispanic Control Added	2.023**	0.088**	0.040**	0.0904**
	[0.133]	[0.010]	[0.010]	[0.010]
Drug Charges Excluded from Charging	2.069**	0.086**	0.049**	0.0835**
	[0.124]	[0.009]	[0.009]	[0.010]
Any Drugs Mentioned in File	2.255**	0.083**	0.052**	0.0795**
	[0.177]	[0.009]	[0.009]	[0.010]
Marital Status	2.001**	0.086**	0.058**	0.0842**
	[0.120]	[0.010]	[0.010]	[0.010]
Exclude FBI	1.894**	0.088**	0.047**	0.0927**
	[0.113]	[0.010]	[0.009]	[0.011]
Exclude ATF	2.175**	0.090**	0.065**	0.0701**
	[0.157]	[0.011]	[0.011]	[0.012]
Property Crimes & Regulatory Crimes	2.32**	0.078**	0.045**	0.0189
	[0.311]	[0.016]	[0.015]	[0.018]
Violent and Gun Crimes	1.930**	0.103**	0.063**	0.123**
	[0.125]	[0.011]	[0.012]	[0.010]

Each cell contains the coefficient on the black dummy from a separate regression. Standard errors clustered by district-arrest offense are in brackets. * p<0.05, **p<0.01

Table A3: Conviction Severity [Post-Plea Bargaining]

	[1] Mand. Min Dummy [Odds Ratio]	[2] Log Stat Max OLS	[3] Log Guideline Sentence OLS	[4] Log Past Mean Sentence OLS
Black	1.933**	0.0732**	0.0234*	0.0647**
	[0.131]	[0.0103]	[0.0112]	[0.0101]
Age	0.972**	-0.00335**	-0.00217**	-0.00356**
	[0.003]	[0.0005]	[0.0005]	[0.0005]
Multi-Defendant	2.092**	0.123**	-0.0414*	0.164**
	[0.153]	[0.0154]	[0.0198]	[0.0164]
Poverty Rate	0.988	0.00296	0.00484	0.00297
	[0.012]	[0.0024]	[0.0026]	[0.0029]
Unemployment Rate	1.012	0.00311	-0.00554	-0.00503
	[0.031]	[0.0073]	[0.0106]	[0.0078]
Log Per Capita Income	0.656	0.0984	0.00199	0.0872
	[0.159]	[0.0558]	[0.0497]	[0.0522]
Crime Rate	1.004	-0.00018	-0.0018	0.000473
	[0.012]	[0.0023]	[0.0023]	[2.3886]
Observations	30325	29686	29578	24853

Col. 1 gives odds ratios from a logistic regression; Cols. 2-4 contain OLS coefficients. Outcome variables are based on the combined final charges of conviction. The sample consists of cases in which the *conviction* took place in fiscal years 2007-09. All regressions include district and arrest-offense fixed effects. Standard errors clustered by offense-district. * $p<0.05$, **$p<0.01$

Table A4: Sentencing Disparity Explained by Mandatory Minimum Charging, Additional Robustness Checks

	Mean	10	20	30	40	50	60	70	80	90	Observations
Main Specification [Table 5b]	0.068** [0.007]	0.047** [0.006]	0.047** [0.005]	0.055** [0.006]	0.061** [0.006]	0.067** [0.007]	0.073** [0.008]	0.084** [0.009]	0.090** [0.010]	0.075** [0.009]	25695
Original Arrest Offense	0.067** [0.007]	0.047** [0.006]	0.046** [0.005]	0.054** [0.006]	0.061** [0.006]	0.066** [0.007]	0.073** [0.008]	0.083** [0.009]	0.090** [0.010]	0.073** [0.009]	25860
Include Hispanic Flag	0.069** [0.007]	0.049** [0.006]	0.049** [0.006]	0.055** [0.006]	0.061** [0.006]	0.067** [0.007]	0.074** [0.008]	0.086** [0.009]	0.090** [0.010]	0.075** [0.009]	22799
Drug Charges Excluded from Charging	0.068** [0.007]	0.047** [0.006]	0.047** [0.005]	0.054** [0.006]	0.061** [0.007]	0.067** [0.007]	0.074** [0.008]	0.084** [0.009]	0.091** [0.010]	0.076** [0.009]	25695
Any Drugs Mentioned in File	0.063** [0.006]	0.045** [0.006]	0.044** [0.005]	0.050** [0.005]	0.055** [0.006]	0.060** [0.006]	0.066** [0.007]	0.075** [0.008]	0.081** [0.009]	0.070** [0.008]	25695
Exclude ATF Arrests	0.080** [0.008]	0.056** [0.008]	0.058** [0.007]	0.059** [0.006]	0.070** [0.007]	0.079** [0.008]	0.088** [0.009]	0.096** [0.010]	0.103** [0.012]	0.097** [0.012]	18550
Exclude FBI Arrests	0.052** [0.006]	0.046** [0.006]	0.038** [0.005]	0.044** [0.006]	0.049** [0.006]	0.053** [0.007]	0.056** [0.007]	0.063** [0.008]	0.066** [0.009]	0.055** [0.008]	21113
Property Crimes	0.055** [0.011]	0.028** [0.007]	0.027** [0.006]	0.035** [0.007]	0.051** [0.010]	0.060** [0.012]	0.061** [0.012]	0.061** [0.012]	0.069** [0.014]	0.057** [0.015]	8911
Gun and Violent Crimes	0.061** [0.008]	0.043** [0.006]	0.047** [0.006]	0.056** [0.008]	0.063** [0.008]	0.066** [0.009]	0.068** [0.009]	0.073** [0.010]	0.072** [0.010]	0.047** [0.007]	16679

Each cell is from a separate Oaxaca pooled decomposition of the black-white gap at the mean and in the RIF at the deciles. Each cell contains the percentage points of the black-white sentence gap that can be explained by the presence of a charge carrying a mandatory minimum sentence. All decompositions include controls for arrest offense, criminal history, education, district fixed effects, multiple defendant flag, age, and county unemployment, poverty, and crime rates as well as per capita income. Standard errors clustered by arrest-offense-district are in brackets. * $p<0.05$, **$p<0.01$

Table A5: Persisting Unexplained Racial Variation in Sentences, Alternate Specifications and Samples

|	Mean	10	20	30	40	50	60	70	80	90	Observations
Main Specification (Table 5b)	0.048** [0.016]	-0.085* [0.039]	0.102** [0.024]	0.011 [0.021]	0.073** [0.020]	0.077** [0.017]	0.021 [0.019]	0.009 [0.019]	0.067** [0.023]	0.146** [0.028]	25695
Exclude Pre-Arrest Indictments	0.069** [0.019]	-0.063 [0.046]	0.129** [0.031]	0.075** [0.026]	0.083** [0.025]	0.079** [0.022]	-0.002 [0.022]	0.038 [0.023]	-0.029 [0.029]	0.209** [0.030]	14696
South Only	0.045* [0.022]	-0.078 [0.056]	0.014 [0.039]	0.029 [0.033]	-0.000 [0.030]	0.075** [0.026]	0.033 [0.026]	0.147** [0.025]	0.089** [0.031]	-0.013 [0.038]	13067
High Crime Counties	0.017 [0.021]	-0.230** [0.064]	0.073* [0.033]	0.028 [0.028]	-0.008 [0.025]	0.030 [0.023]	-0.020 [0.023]	-0.012 [0.027]	-0.035 [0.030]	0.097** [0.035]	13493
Counsel Type Recorded	0.089** [0.023]	-0.102 [0.070]	0.169** [0.041]	0.077* [0.034]	0.068* [0.031]	0.122** [0.027]	0.016 [0.029]	0.006 [0.030]	0.023 [0.038]	-0.020 [0.042]	9900
Includes Counsel Type Control	0.088** [0.023]	-0.108 [0.070]	0.175** [0.041]	0.073* [0.034]	0.069* [0.031]	0.121** [0.027]	0.017 [0.029]	0.008 [0.030]	0.024 [0.038]	-0.020 [0.043]	9900
Police Notes Available	0.046* [0.022]	-0.052 [0.052]	0.034 [0.037]	0.087** [0.032]	0.019 [0.028]	0.081** [0.025]	-0.002 [0.026]	0.027 [0.028]	-0.042 [0.037]	0.006 [0.038]	11265
Includes Police Notes	0.045* [0.022]	-0.054 [0.052]	0.032 [0.037]	0.087** [0.032]	0.019 [0.028]	0.082** [0.025]	-0.001 [0.026]	0.028 [0.028]	-0.042 [0.037]	0.003 [0.039]	11265
Mandatory Minimum Conviction											
Authors' Coding Scheme	0.051** [0.016]	-0.080* [0.039]	0.107** [0.024]	0.016 [0.021]	0.077** [0.019]	0.081** [0.017]	0.025 [0.018]	0.011 [0.019]	0.070** [0.022]	0.145** [0.028]	25695
USSC Coding	0.015 [0.016]	-0.093* [0.039]	0.091** [0.024]	-0.005 [0.021]	0.051** [0.019]	0.049** [0.017]	-0.014 [0.018]	-0.043* [0.018]	-0.000 [0.022]	0.059* [0.028]	25661
USSC Coding: Includes Drugs & Child Pornography Cases	0.064** [0.011]	0.033 [0.028]	0.011 [0.019]	0.027 [0.016]	0.147** [0.013]	-0.011 [0.014]	0.096** [0.014]	0.164** [0.014]	-0.032* [0.014]	0.034 [0.018]	58861

Each cell contains the unexplained disparity from a separate Oaxaca pooled decomposition of the black-white gap at the mean and in the RIF at the deciles (analogous to those in Table 6). All decompositions include controls for arrest offense, criminal history, education, district fixed effects, multiple defendant flag, age, and county unemployment, poverty, and crime rates as well as per capita income. Standard errors clustered by arrest offense-district are in brackets.

Table A6: Persisting Unexplained Racial Variation in Sentences, Additional Robustness Checks

	Mean	10	20	30	40	50	60	70	80	90	Observations
Main Specification [Table 5b]	0.048** [0.016]	-0.085* [0.039]	0.102** [0.024]	0.011 [0.021]	0.073** [0.020]	0.077** [0.017]	0.021 [0.019]	0.009 [0.019]	0.067** [0.023]	0.146** [0.028]	25695
Original Arrest Offense	0.046** [0.016]	-0.090* [0.039]	0.099** [0.024]	0.009 [0.021]	0.071** [0.020]	0.075** [0.017]	0.018 [0.019]	0.006 [0.019]	0.065** [0.023]	0.142** [0.028]	25860
Include Hispanic Flag	0.049** [0.018]	-0.068 [0.044]	0.099** [0.028]	0.035 [0.023]	0.013 [0.021]	0.084** [0.018]	-0.031 [0.020]	0.011 [0.021]	0.105** [0.024]	0.118** [0.031]	22799
Drug Charges Excluded from Charging	0.047** [0.016]	-0.086* [0.039]	0.102** [0.025]	0.011 [0.021]	0.073** [0.020]	0.077** [0.017]	0.020 [0.019]	0.008 [0.019]	0.066** [0.023]	0.144** [0.028]	25695
Any Drugs Mentioned in File	0.046** [0.016]	-0.086* [0.039]	0.101** [0.024]	0.009 [0.021]	0.070** [0.019]	0.074** [0.017]	0.017 [0.018]	0.005 [0.019]	0.062** [0.023]	0.143** [0.028]	25695
Exclude ATF Arrests	0.017 [0.019]	-0.081 [0.053]	0.101** [0.031]	-0.012 [0.024]	-0.020 [0.022]	-0.019 [0.022]	0.070** [0.021]	-0.018 [0.023]	0.100** [0.025]	0.154** [0.031]	18550
Exclude FBI Arrests	0.057** [0.017]	-0.091 [0.051]	0.119** [0.026]	0.033 [0.022]	0.012 [0.022]	0.134** [0.019]	0.025 [0.021]	-0.002 [0.021]	0.109** [0.025]	0.068* [0.030]	21113
Property Crimes	-0.055 [0.029]	-0.019 [0.044]	-0.044 [0.028]	-0.057* [0.029]	-0.064* [0.031]	-0.078** [0.027]	-0.076** [0.027]	-0.099** [0.030]	-0.069* [0.034]	-0.070 [0.045]	8911
Gun and Violent Crimes	0.094** [0.014]	0.204** [0.029]	0.117** [0.022]	0.120** [0.020]	0.127** [0.019]	0.095** [0.018]	0.147** [0.018]	0.102** [0.019]	0.044 [0.024]	-0.059** [0.020]	16679

Each cell is from a separate Oaxaca pooled decompositions of the black-white gap at the mean and in the RIF at the deciles. Each cell contains the percentage points of the black-white sentence gap that can be explained by the presence of a charge carrying a mandatory minimum sentence. All decompositions include controls for: arrest offense, criminal history, education, district fixed effects, multiple defendant flag, age, and county unemployment, poverty, and crime rates as well as per capita income. Standard errors clustered by arrest-offense-district are in brackets. * p<0.05, **p<0.01

Figure A1: Disparity Explained by Mandatory Minimum Charges [percentage points]: Alternate Specifications and Samples [Table 6][1]
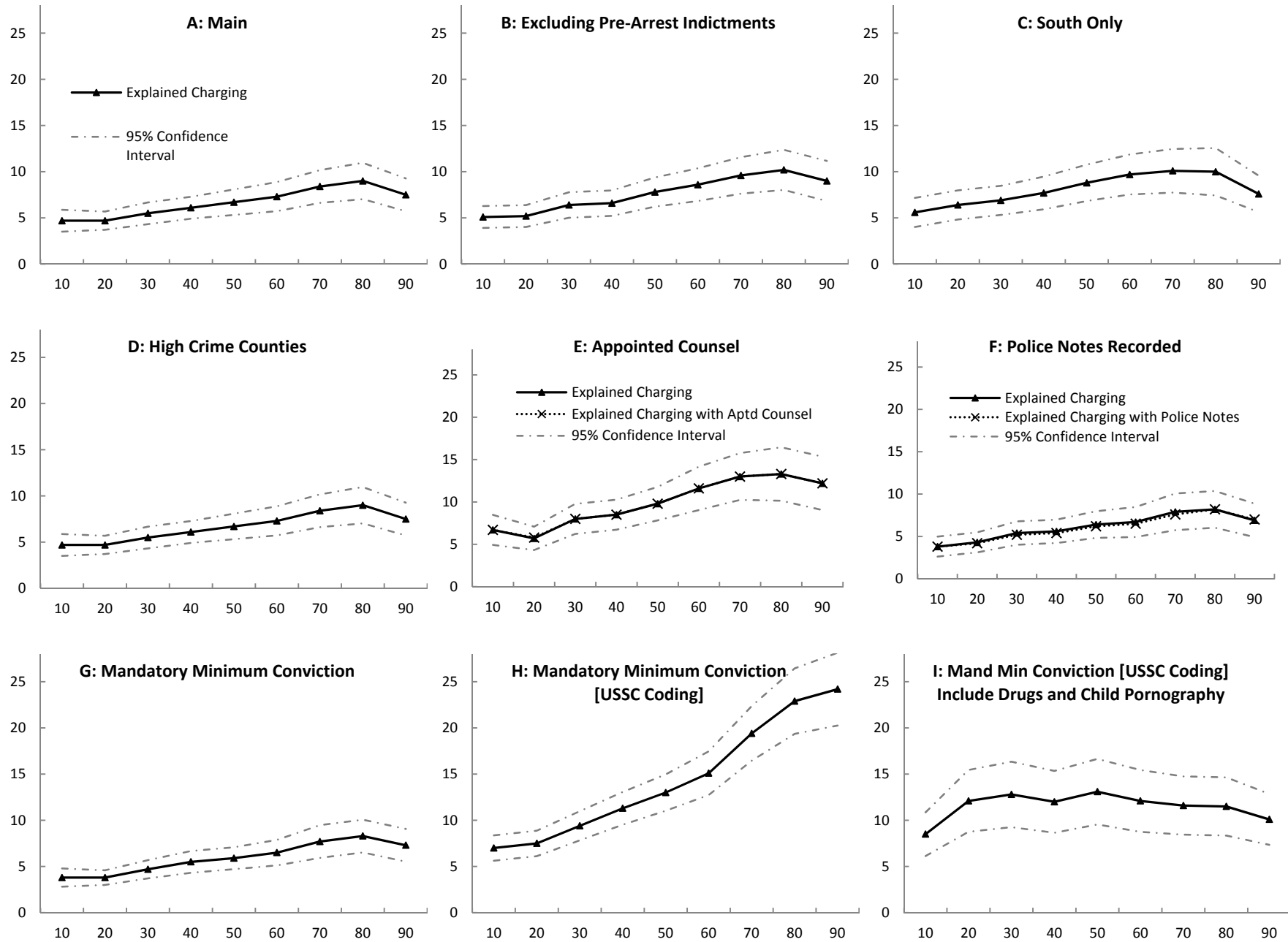
[1] All figures contain point estimates and 95% confidence intervals estimated from a RIF-Oaxaca regression with all of the controls contained in Table 5b.

Figure A2: Unexplained Black-White Disparity in Sentence Length by Decile of the Sentencing Distribution [Tables 5A and 5B]