

RESEARCH REPORT

# Looking Beyond the Sentence

Examining Policy Impacts on Racial Disparities in Federal Sentencing across Stages and Groups and over Time

*Mari McGilton*  
*January 2023*

*Sabrina Rizk*



## ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places.

# Contents

<b>Acknowledgments</b>	<b>iv</b>
<b>Executive Summary</b>	<b>v</b>
Methods	v
Findings and Recommendations	vii
<b>Looking Beyond the Sentence</b>	<b>1</b>
Collateral Effects of Disparities in the Criminal Legal System	1
Racial Disparities in the Criminal Legal System	2
Racial Disparities in Federal Sentencing	3
Targeting Research to Understand Disparities	5
The Federal Sentencing Process	6
Methods	8
Sample	8
Dependent Measures	12
Independent Measures	12
Analytic Strategy	13
Subgroup Analyses	13
Policy Review	14
Results	14
Disparities across Stages and over Time	15
Policy Review	23
Key Findings and Recommendations	30
Limitations	32
<b>Notes</b>	<b>35</b>
<b>References</b>	<b>36</b>
<b>About the Authors</b>	<b>41</b>
<b>Statement of Independence</b>	<b>42</b>

# Acknowledgments

This report was funded by the National Institute of Justice (15PNIJ-21-GG-02817-RESS). We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute’s funding principles is available at [urban.org/fundingprinciples](https://urban.org/fundingprinciples).

We thank Nancy Merritt for her ongoing support. We also thank former and current Urban Institute staff for their assistance with this project: Bill Adams, Ashlin Oglesby-Neal, and Olivia Stuart. Our special thanks to David Pitts and Zach VeShancey for their thorough review and constructive input on drafts of this report.

# Executive Summary

Pressure to identify and reduce disparities through policy has risen in recent years, but these priorities are neither new nor easily achieved. Most studies of these disparities are siloed from each other and are limited to a single outcome measure, comparison group, and/or comparison time point, meaning policymakers end up relying on incomplete information to make critical decisions. This is especially true for federal and state sentencing decisionmakers, who are bound by policies made across multiple agencies. Federal sentencing decisionmakers specifically are driven by US attorney general directives, legislative policies, Supreme Court rulings, and amendments to the US federal sentencing guidelines. These policies affect several decisionmakers at different stages in the federal sentencing process (e.g., attorneys' application of mandatory minimum statutes, federal probation officers' determination of final offense level, federal judges' sentence decisions). Most studies on disparities at the federal sentencing stage of the criminal legal system investigate only racial disparities in sentencing decisions, such as in/out decisions (i.e., prison/no prison), and/or sentence lengths, and evaluate changes in these disparities based on predetermined points in time from predetermined policies assumed to have impacted those changes. Although these studies provide critical information on disparities in federal sentencing, we could expand our knowledge and the methods used.

Funded by the National Institute of Justice, we sought to answer three key research questions:

1. Do racial disparities vary across three stages of federal sentencing and over time? If so, how?
2. During which years do the measured racial disparities have a statistically significant decrease?
3. Which policies likely impacted these decreases the most? What are the commonalities between them?

## Methods

We used publicly available federal sentencing data for FY1999 through FY2021, which consisted of 1,623,744 cases. After obtaining the dataset, we created dichotomous variables for the four major race groups in the United States Sentencing Commission (USSC) data—(1) white, (2) Black, (3) American Indian and Alaska Native (AIAN), and (4) Asian and Pacific Islander (API)—to be the “treatment-comparison” variables. For example, white people are in the treatment group and nonwhite people are in the matched comparison group. We did not include Hispanic origin as a separate racial/ethnic group and instead used it as a covariate.<sup>1</sup> After creating the dichotomous race variables, we used exact

matching on age group, sex, primary sentencing guideline, district of sentencing, fiscal year, education status, whether the person was of Hispanic origin, US citizenship status, and whether the case was resolved through a trial or guilty plea.

We measured racial disparities across three stages of federal sentencing, each of which is represented by two key elements (for a total of six elements): (1) the charge stage, represented by mandatory-minimum eligibility and statutory maximum eligible sentence; (2) the computation stage, represented by the final offense level and total criminal history points; and (3) the sentence stage, represented by total sentence length and the number of months sentenced below the maximum of the guideline range. We used separate ordinary least squares and logistic regressions to measure the disparities between the matched white and nonwhite groups for each stage. We did this for each fiscal year, then gathered the coefficients for each fiscal year into a time-series object. Once we created time-series objects for each stage, we detected structural breaks (i.e., breakpoints) in the trend of differences for each stage across fiscal years.

To deepen our understanding of racial disparities at different stages of sentencing and over time, we conducted a series of subgroup analyses. In addition to examining differences in disparities between stages of federal sentencing, we sought to understand differences in disparities between racial groups. Thus, we changed the reference group (i.e., white) to the other three racial groups identified in the USSC data—Black, AIAN, and API—and remeasured the disparities to examine whether the originally measured differences are consistent between groups and over time. After measuring the size of disparities for each matched group over time, we compared the disparities of each matched group to one another to examine differences between those groups. To examine whether the measured disparities differed by offense type, we compared the disparities between white and nonwhite groups for the four most common guidelines in Chapter 2 of the US federal sentencing guidelines. Consistent with the main analysis, we collected the logged coefficients and odds ratios for each subgroup into time-series objects and detected structural breaks across fiscal years.

Using the breakpoints, we reviewed policies that might have contributed to significant decreases in racial disparities across the six elements of federal sentencing. We searched all legislation, Supreme Court decisions, amendments to the US federal sentencing guidelines, and US attorney general directives enacted during the year of each breakpoint, and up to two years earlier, to identify which likely contributed to the significant decreases in racial disparities.

## Findings and Recommendations

Racial disparities differ at each key stage and for each element of federal sentencing. We consistently found the largest racial disparities for matched comparisons and the greatest differences between subgroups in the charge stage. Total criminal history points also had quite a bit of variability, high racial disparities, and large differences between groups. Conversely, total sentence length consistently had the smallest measured disparities over time and the smallest differences between groups. This is especially important because this is the element used most often in racial disparity studies. Although critical to understand, conclusions drawn about the size and nature of racial disparities in federal sentencing have likely been understated because sentence length and type have been the sole focus of such studies. With that, we only measured one breakpoint (the same breakpoint across matched racial groups) for the total sentence length element, which may mislead interpretations of policy impacts if it continues to be the only area of focus in data and policy studies. Given the differences between elements and groups, **we recommend that policymakers and researchers consider the entire federal sentencing process together when investigating disparities.**

Each breakpoint we examined indicated a fiscal year when a significant decrease in disparities between matched groups occurred. The periods with the greatest decreases in racial disparities were similar across elements and generally occurred in FY2004/FY2005, FY2012 to FY2014, and FY2016 to FY2018. Although there are similarities between the breakpoints across elements, there are differences in the breakpoints between the white/nonwhite and Black/non-Black disparities. The sentence length element is the exception, which supports the need to explore disparities in additional outcomes and groups.

To expand our knowledge beyond the traditional white/nonwhite comparisons, we measured disparities across elements and over time between matched white and nonwhite, Black and non-Black, AIAN and non-AIAN, and API and non-API people. We found that the sizes, directions, and trends of disparities differed greatly between groups. For example, Black people are most disadvantaged across all elements and periods (e.g., greater number of criminal history points, longer sentence lengths), whereas AIAN people had the highest advantages (e.g., lower eligible statutory maximums, lower final offense levels). Disparities were largest between Black and non-Black people, whereas disparities were consistently smallest between API and non-API people and white and nonwhite people. The experiences of people of different races throughout the federal sentencing process differ, and the breakpoints indicate that policies impact them differently. Given the stark differences in the disparities between racial groups, the differences between stages in the sentencing process, and different trends over time, **we recommend measuring and accounting for disparities among all demographic groups, in**

**addition to those between white and nonwhite or white and Black people, to understand the broad range of experiences of all people in the sentencing process.**

The differences between the disparities for each primary sentencing guideline demonstrate that racial disparities differ by the offense under which people are sentenced. Many policies target specific crimes and thus can impact racial disparities for those crimes individually. But the data demonstrated many similarities between the primary sentencing guidelines, indicating that although certain policies may not be intended for specific crimes, they may still be impacted. For example, drug trafficking offenses constitute the largest share of federal sentences and likely influenced the changes in racial disparities overall. In our policy analysis, we found that within two years of the enactment of each policy targeting drug offenses (e.g., the Fair Sentencing Act, the Crack Minus Two Amendment, and the Drugs Minus Two Amendment), overall disparities decreased significantly. But the breakpoints specific to drug trafficking offenses differed from the general breakpoints. The data show that sentencing policies related to drug trafficking clearly impacted cases other than those they were intended to impact, and their effects overlap with those of other policies. Therefore, **we recommend that policymakers and researchers extend impact analyses beyond the apparently relevant policies, offense types, and groups to fully understand the breadth and depth of policy impacts.**

Policies impact groups differently, reach across stages in the sentencing process, and impact multiple system actors. We saw this often in the charge stage, which consistently had the largest disparities over time and between groups. After each US attorney general issued a directive, a breakpoint occurred. Although each directive was intended to standardize prosecution and plea deal practices, they differed in approach. Each directive lacked instructions for how it was to be carried out, which likely explains the large variability in the disparities over time. Because of the apparent impact the US attorney general has on disparities, **we recommend further research into the effects the US attorney general and US attorneys have on racial disparities throughout the sentencing process.**

The importance of clear and operational policies also became apparent in our analysis. It became evident during our examination that many breakpoints were likely the combined impact of policies. Many policies identified as likely contributors to the breakpoints in disparities were enacted years before the effect was registered. During those periods, multiple policies were enacted to clarify and operationalize the original policies and likely combined to produce the registered effect. For example, the Fair Sentencing Act had a clear impact on racial disparities, particularly for Black people, but it took several years to register a significant impact. Its impact was likely spread out over time because several subsequent policies had to be passed to interpret and standardize its implementation (e.g., USSC amendment 750, *Dorsey v. United States*). This may have also impacted the implementation of the



directive issued by US Attorney General Holder in 2010. Inconsistent and vague language was the most common reason for the amendments to the US federal sentencing guidelines relating to breakpoints for total criminal history points and is likely why disparities vary so greatly over time compared with the other computation element, the final offense level. Conversely, policies such as the 2001 amendments to the US federal sentencing guidelines, followed by the enactment of the Sarbanes-Oxley Act<sup>2</sup> and the accompanying US attorney general directive, all provided detailed information on offenses and how to charge and sentence, and likely had an immediate impact beginning in FY2002. Future studies may find something similar when evaluating the impact of the First Step Act. With clear, specific, and operational language, the First Step Act was likely enacted by all system actors at all stages uniformly. In the nearly four years since its passing, no additional breakpoints have registered, and no clarifying policies have been enacted. Thus, **we recommend that policymakers include clear statements in their laws, directives, rulings, and amendments that are operational from the outset and require minimal clarification or subjective interpretation.**

Retroactive policies and social factors likely impacted the racial disparities measured in this study as well. For instance, in the later years (FY2016 and later), the disparities flipped, as those who typically experienced advantages suddenly began experiencing disadvantages, although no breakpoint registered. From contentious presidential elections, increased conversations and awareness about racial injustices, and the later outbreak of COVID-19, any significant events in the past five years could explain the flip seen in this study. Although descriptive studies have been done on retroactive policies (e.g., Hunt, Rutter, and Kostyshak 2020; USSC 2014a, 2021), **we recommend further inquiry into the impact of retroactive policies and social factors on disparities, and additional examination of the aforementioned flip in a few years to understand the long-term change and its contributing factors.**



# Looking Beyond the Sentence

Racial disparities in the criminal legal system have a profound and far-reaching impact on individuals, families, communities, and the economy (Couloute and Kopf 2018). Pressure to identify and reduce disparities through policy has risen in recent years; but these priorities are neither new nor easily achieved. Identifying and reducing racial disparities at all points in the criminal legal system is crucial, given the devastating consequences system involvement can have on people, families, and communities. Most studies of these disparities are siloed from each other and are limited to a single outcome measure, comparison group, and/or comparison time point, meaning policymakers end up relying on incomplete information to make critical decisions. This is especially true for federal and state sentencing decisionmakers, who are driven by mandates made in multiple agencies. Federal sentencing decisionmakers have to follow mandates by US attorney general directives, legislative policies, Supreme Court rulings, and amendments to the US federal sentencing guidelines. These mandates affect multiple decisionmakers at multiple decision points in the sentencing process. Racial disparities are not solely a criminal legal system issue given the impact criminal legal system involvement has on individuals, families, and communities. It is imperative to understand the historical and present racial disparities in the criminal legal system to eliminate them, thereby reducing adverse effects on communities, especially communities of color. With funding from National Institute of Justice, we **measured racial disparities across stages** in the federal sentencing process **for multiple racial groups**, **detecting when statistically significant changes occurred** in measured racial disparities between fiscal years 1999 and 2021, and **identified which policies likely impacted significant decreases most** (e.g., legislation, US attorney general directives, Supreme Court rulings, amendments to the US federal sentencing guidelines). In this report we discuss the purpose of this study and our methods, results, and conclusions.

## Collateral Effects of Disparities in the Criminal Legal System

Criminal legal system involvement is a lifelong burden, even if it does not involve a life sentence. Once someone is arrested, whether they are convicted or not, a record is created that will flag them as “justice involved.” A longitudinal survey conducted by RAND Corporation on the rates and outcomes of arrests and convictions of Americans found that arrests are rising significantly across the country (Smith 2018). Arrests early in life lead to adverse outcomes including lower overall income, lower

marriage rates, and lower educational attainment (Smith 2018). Once convicted, a person endures punishment for the crime for the remainder of their life. A felony conviction creates barriers to housing, loans, employment, and social services (Pinard 2010; Smith 2018; Wheelock 2005). Research has found that all these outcomes disproportionately affect Black people and other marginalized populations (Administration for Children and Families 2000; Fellner and Mauer 1998; Uggen and Manza 2002; Wheelock 2005). Employers are approximately 50 percent less likely to call back Black job applicants with criminal records than Black applicants without criminal records (Bushway 1996; Dietrich 2002; Holzer, Raphael, and Stoll 2001; Pager 2003). This is particularly troublesome given that people who are Black and do not have criminal records receive fewer callbacks than people who are white and have criminal records (Pager 2003).

Not only do people with conviction histories face additional consequences after their sentences, but their families do too. Family members can experience exclusion and stigmatization from their communities, financial stress, disrupted emotional attachments, and loss of parental figures (Arditti 2012; Condry and Minson 2020). If the person who was arrested and sentenced was their family's sole provider and caregiver, their family could fall into poverty and their children could be removed from their home and placed in foster care (Condry, Kotova, and Minson 2016).

Given the different impacts criminal legal system involvement has on individuals, families, and communities, racial disparities in criminal legal system involvement are not solely a criminal legal system issue. It is imperative that we understand the historical and present racial disparities in that system so we can eliminate them and reduce their adverse effects on communities, especially communities of color.

## Racial Disparities in the Criminal Legal System

Disparities in the US criminal legal system are rooted in the country's history and date to the country's inception, with the first anti-immigration laws and the exile and imprisonment of Mexican and Indigenous populations to acquire land and resources (Anderson 2005; Debrah 2012; Fischer 2015; LeBrón 2019). Then from slavery through Reconstruction and the present, explicit and implicit policies have driven disparities in the criminal legal system that have disproportionately impacted nonwhite communities. The actions meant to maintain slavery after the passing of the Thirteenth Amendment set the foundation for racial disparities in the criminal legal system today. During Reconstruction, policy was used to control people of color through the criminal legal system (Hinton, Henderson, and Reed 2018). Policies including the Black Codes, vagrancy laws, and convict leasing directly targeted people of

color (especially Black people) by systematically reincorporating forced labor and confinement as modes of punishment for petty crimes (Blackmon 2008; Haley 2017). After the civil rights movement, policies moved away from overt forms of racism toward more covert forms, driven by myths about people of color inherently needing to commit crimes (Headley 1983). The inception of criminal legal policies through the war on crime, the war on drugs, the Violent Crime Control and Law Enforcement Act, and immigration laws, and the embracing of those policies by policymakers, increased prosecution and penalties that disproportionately affected nonwhite communities (Kohler-Hausmann 2010; Moore and Elkavich 2008; Thompson 2019). Statistics and research have been used to drive perspectives that people of color commit more crimes and are more dangerous, therefore promoting increased intervention by police and the criminal legal system in nonwhite neighborhoods (Headley 1983; Hinton 2016; Muhammad 2010). The proportion of people of color in prison for any crime consistently overrepresents the actual population of color in the United States (Hinton, Henderson, and Reed 2018; Sentencing Project 2008; Schrantz and McElroy 2018). However, these findings and statistics rely on the notion that policies are applied uniformly, disregarding the history of our criminal legal system and the reasoning behind our current policies.

## Racial Disparities in Federal Sentencing

Efforts to end disparities in sentencing, especially federal sentencing, date back nearly 50 years. Until the mid-1970s, indeterminate sentencing was dominant for judges and parole boards, with the focus on punishment being individual rehabilitation. Judges and parole boards had the discretion to increase or decrease punishment on a case-by-case basis, which advocates and scholars eventually criticized for driving disparities in the criminal legal system. From the mid-1970s to the mid-1980s, policies such as Truth in Sentencing and the passage of the Sentencing Reform Act created a structure in sentencing and predictability in punishment and resulted in the establishment of the US Sentencing Commission (USSC). From the mid-1980s through the mid-1990s, the combination of structured sentencing and the inception of harsher punishments, such as the three-strikes law and mandatory minimum penalties for certain drugs and weapons offenses, caused federal sentencing lengths and prison populations to skyrocket (USSC 2017, 2018b). People who had longer criminal histories because of other policies, such as broken windows policies, hot-spot policing, and stop, question, and frisk, were at higher risk of receiving overly harsh sentences and were mostly people of color (ACLU 2014; Hinton, Henderson, and Reed 2018; NASEM 2018; Rehavi and Starr 2014). From the mid-1990s to the mid-2000s, policies aimed at increased punishments for sex crimes and child pornography crimes were prevalent (e.g., the Adam Walsh Child Protection and Safety Act and Megan's Law), though their impacts on racial

disparities appear to be less than those of other policies. From the mid-2000s until now, most federal policies have aimed at lessening punishments and reducing disparities (e.g., the Fair Sentencing Act and the First Step Act).<sup>3</sup> In addition, several Supreme Court rulings (e.g., *United States v. Booker* and *Gall v. United States*)<sup>4</sup> and amendments to the US federal sentencing guidelines (e.g., drug/marijuana equivalency recalculations, first-time offender considerations, Safety Valve, and the Drugs Minus Two Amendment) have aimed at reducing disparities in punishments by standardizing the sentencing process (USSC 2019a).

In addition to policies, racial disparities result from how key decisionmakers exercise discretion. From prosecutorial discretion in charging, plea bargaining, and/or the application of mandatory minimum penalties, to judicial discretion over sentence type and length, researchers have found that people of color are consistently treated more harshly than white people at every stage of sentencing (Kutateladze, Lynn, and Liang 2012; Rachlinski et al. 2007; Rehavi and Starr 2014; Schmitt, Reedt, and Blackwell 2017). As the gatekeepers to the criminal legal system, prosecutors hold the power to decide whether to charge people and with what, impacting the rest of the process through sentencing and collateral consequences (Spohn 2013; Vera 2017).<sup>5</sup> Kutateladze, Lynn, and Liang (2012) found that before deciding whether and what to charge, prosecutors consider the strength of the evidence, seriousness of the offense, characteristics of the victims and whether they want and are willing to testify, the idea of doing justice and returning balance to the community, and the defendant's legal and personal characteristics and circumstances. They also found that prosecutors are often restricted by (1) internal rules or policies regarding specific cases or (2) a lack of resources that might motivate them to dismiss charges deemed relatively less serious. Typically, decisionmakers in the court system work in the same groups and on similar cases, becoming familiar with one another's preferences and work styles. These relationships can influence their decisions. For example, suppose a prosecutor knows about the judge's track record of hearing or rejecting cases. They may only file charges they know will be accepted or dismiss cases that will be rejected (Frederick and Stemen 2012). Federal prosecutors specifically have a great deal of power to use their discretion in charging and conviction practices, though their discretion is heavily influenced by US attorney general directives.

The discretion of federal probation officers on what information to include in the presentencing reports used during sentencing and how to do so primarily guides judges in handing down sentences. Little is known about the impact of federal probation officers' discretion in constructing presentence reports, however. As the final word regarding sentence type and length for people convicted of crimes, the discretion of federal judges is paramount in people's lives after sentencing, which is likely why sentence in/out decisions and sentence length are the focal points of most research on racial disparities

in federal sentencing. In recent studies about the sentencing practices of federal judges, the USSC (2019b, 2020b) found that individual differences exist between the sentencing practices of judges in the same cities and in different cities. These outcomes are true even when controlling for other factors, such as criminal history (Schmitt, Reedt, and Blackwell 2017). With that, an analysis of 40 studies of racial disparities in federal and state sentencing found that 43 to 68 percent had “racially discriminatory” sentencing outcomes, including in/out imprisonment decisions and overall sentence length (Kansal 2005).

Between legislation, Supreme Court rulings, amendments, and the individual discretion of critical decisionmakers, multiple layers of compounding factors contribute to these racial disparities in federal sentencing, making it difficult to determine with certainty that any one factor has resulted in a change in those disparities. Efforts to understand and answer (even partially) the degree to which these factors impact racial disparities are imperative for creating policies to eliminate disparities in the criminal legal system.

## Targeting Research to Understand Disparities

Most research on racial disparities in federal sentencing focuses on identifying disparities related to sentence length or prison in/out outcomes (Franklin and Henry 2020; Mitchell and MacKenzie 2004; Mustard 2001; Nowacki 2016; Valadez and Wang 2017). Racial disparities in sentencing may extend far beyond sentence length and type, especially in federal sentencing, which can influence other decisions in the process, such as sentence length or in/out decisions (Feldmeyer and Ulmer 2011; Fischman and Schanzenbach 2012; Tuttle 2018; Yang 2015). For instance, Rehavi and Starr (2014) found that prosecutors’ charging decisions (i.e., pursuing statutes that enact mandatory minimums over lesser charges) explained more than half of unexplained racial disparities. Further, Kim, Spohn, and Hedberg (2015) used cross-classified random effects models to examine the judges’ and prosecutors’ roles in sentences, individually and collectively. They found that sentences varied by over a year based on the prosecutor assigned the case. Conversely, sentences varied by six months based on the judge assigned the case. It is imperative to examine multiple stages in federal sentencing in tandem to identify the existence and extent of disparities and the connections between them.

Most studies use white males as the comparison group and only explore disparities between white and Black or white and nonwhite people. In fact, Mitchell and MacKenzie (2004) found that all 85 articles included in their meta-analysis on racial disparities in sentencing used white males as the only investigated group. With that, most research on racial disparities in the criminal legal system examines

one or two predetermined policies at a time (if that), only identifies whether disparities exist, or measures disparities using percentage difference from a sample mean.

Furthermore, policy impacts are usually measured using predetermined points in time and policies (e.g., before/after *United States v. Booker*) (Feldmeyer and Ulmer 2011; Fischman and Schanzenbach 2012; Schmitt, Reedt, and Blackwell 2017; Yang 2015). This approach involves assuming which policies and which points in time may have created change, then calculating the existence and extent of disparities accordingly. In addition, many studies measure disparities generally and only briefly assess policies in their findings and discussions (Johnson and Lee 2013).

Despite the large body of research evaluating racial disparities in the criminal legal system and efforts to reduce them, disparities still exist and harm people and communities. Innovative approaches to identifying, measuring, and acting to eliminate racial disparities are needed. A powerful way to ensure key decisionmakers can make complete, data-driven policy decisions is to measure racial disparities across multiple federal sentencing stages, using matched comparisons, then determine the points in time where statistically significant changes have occurred to target reviews of policies that likely had an impact on these changes (Ferguson 2015; Le Cook et al. 2010; Treuhaft et al. 2020).

We used publicly available USSC data from fiscal years 1999 through 2021 to study racial disparities in federal sentencing. The study consisted of (1) measuring racial disparities at multiple stages in the federal sentencing process, using matched comparisons; (2) using time-series breakpoint analyses to examine statistically significant changes in disparities over time; and (3) systematically examining legislation, directives, Supreme Court rulings, and amendments to the US federal sentencing guidelines in the periods when significant decreases occurred to identify possible drivers of changes in racial disparities.

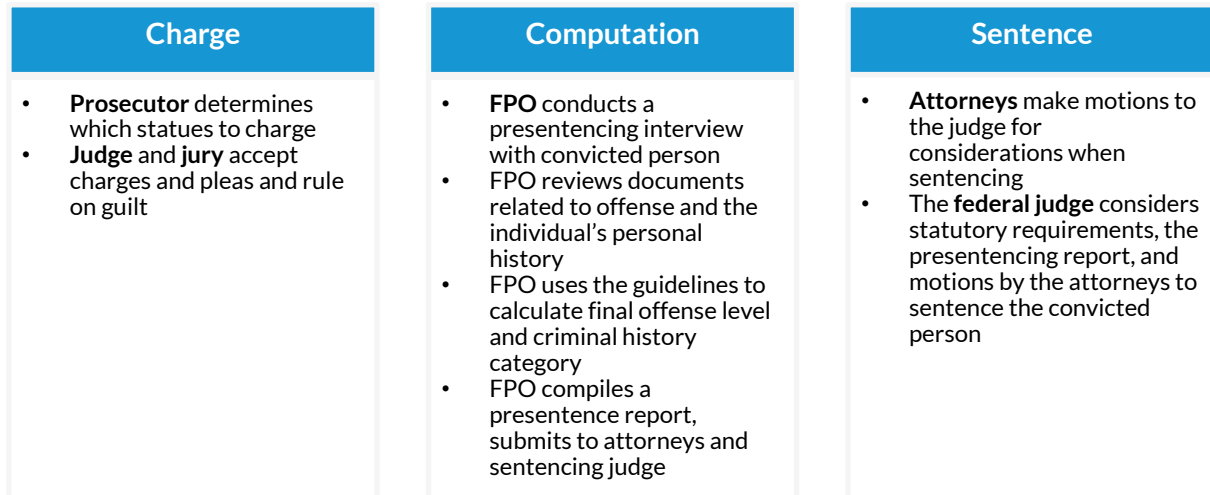
## The Federal Sentencing Process

To identify and evaluate disparities in federal sentencing, it is important to understand how the process works. Federal sentencing begins with the charging and conviction stage (figure 1). After an indictment, the prosecutors determine which statutes to charge a defendant with. The defendant is declared innocent or guilty by a jury's verdict after a trial, or they can choose to submit a plea. Plea bargains are made with the prosecution and contain agreements about how the US federal sentencing guidelines may be applied in a defendant's case, typically to secure charges less than what the prosecution would pursue at trial. Judges cannot participate in plea discussions but are responsible for accepting or rejecting plea agreements.



FIGURE 1

The Federal Sentencing Process



Note: FPO = federal probation officer.

After the judge and/or jury have accepted the charges and pleas and have ruled on guilt, the computation stage begins. A federal probation officer first conducts a presentence interview with the defendant. During the interview, the officer asks questions about the defendant's criminal history, personal and family history, financial circumstances, and any other issues they deem relevant to the court. They will also investigate the defendant's background and criminal record in conjunction with the interview. Using the information they gather and the statutes that are applicable given the charges or pleas the court has accepted, the officer uses the US federal sentencing guidelines manual to calculate the final offense level (offense severity) and criminal history category (severity of past convictions) to determine a recommended sentence range. All the information is compiled into a presentence report that is submitted to the defense and prosecution 35 days before sentencing. The attorneys must submit any objections to the presentence report within 14 days of the sentencing hearing.

During the sentencing hearing, the judge considers the statutory requirements, information provided in the presentence report, motions put forth by the attorneys, and statements made by the person to be sentenced and the victim(s) to rule on a sentence. After the judge pronounces a sentence, the "judgment in a criminal case" document is completed. This document memorializes the judge's oral statement from the hearing and specifies the sentence, including the type and length of sentence, supervision and treatment requirements, and any financial penalties (USSC 2020a).

## Methods

Through this study, we sought to inform the field of observed racial disparities in federal sentencing and how those observed disparities could be reduced or eliminated through policy to promote a fairer system.

We sought to answer three key research questions:

1. Do racial disparities vary across three stages of federal sentencing and over time? If so, how?
2. During which years do the measured racial disparities have a statistically significant decrease?
3. Which policies likely impacted these decreases the most? What are the commonalities between them?

## Sample

To answer our research questions, we used publicly available federal sentencing data for FY1999 through FY2021, which consisted of 1,623,744 cases. We collected data for FY1999 through FY2001 from the National Archive of Criminal Justice Data website and data for FY2002 through FY2021 from the USSC website. We kept only cases that the information on sentencing documents matched (1,429,810 cases). We removed cases where the maximum possible sentence was death. We transformed variables only when it would not introduce or mask racial disparities and was necessary. For example, researchers typically limit the final offense level to 43 because it is the highest value in the federal sentencing table. They typically limit sentence length to 470 months because this equals life in prison. We did not limit the final offense level or sentence length because more people who are Black had final offense levels above 43, and more people who are white had sentences longer than 470 months. We did limit the maximum number of months in the guideline range to 470 because “life” is the maximum allotted value in the US federal sentencing guidelines, defined as 470 months by the USSC. After creating and transforming necessary variables and removing cases with missing or illogical values that could not be confidently inferred, the final dataset contained 1,281,732 observations (79 percent of the original cases; table 1).

TABLE 1

## The Final Sample of People Sentenced Federally between FY1999 and FY2021 Used in This Study

N = 1,281,732

Race		Guideline		Plea or trial	
White	72%	§2D1.1	34%	Plea	96%
Black	24%	§2B1.1	10%	Trial	4%
American Indian / Alaska Native	2%	§2K2.1	9%	<b>Criminal history category</b>	
Asian / Pacific Islander	2%	§2L1.2	21%	Category I	46%
All other racial identities	<1%	§Other	25%	Category II	13%
<b>Hispanic origin</b>				Category III	16%
Non-Hispanic	52%	<b>Circuit</b>		Category IV	10%
Hispanic	48%	DC Circuit	1%	Category V	6%
<b>Sex</b>				Category VI	10%
Male	87%	1st Circuit	3%	<b>Final offense level</b>	
Female	13%	2nd Circuit	6%	1-10	24%
<b>Citizenship status</b>				11-20	32%
US citizen	62%	3rd Circuit	4%	21-30	33%
Non-US citizen	38%	4th Circuit	8%	31-40	11%
<b>Education</b>				> 40	1%
Less than high school	49%	5th Circuit	24%		
High school or more	51%	6th Circuit	8%		
<b>Age group</b>					
<21 years	4%	7th Circuit	4%		
21-25	15%	8th Circuit	7%		
26-30	19%	9th Circuit	18%		
31-35	18%	10th Circuit	8%		
36-40	15%	11th Circuit	10%		
41-50	19%				
>50 years	11%				

Source: Authors' analysis of US Sentencing Commission individual sentencing data, FY1999 through FY2021.

We used RStudio's *Matchit* package to perform exact matching on the data.<sup>6</sup> We used exact matching to limit as many legal and extralegal factors that contribute to federal sentencing decisions as possible to understand the extent to which race plays a role in federal sentencing outcomes on its own. With exact matching, each combination of covariate level creates a subclass for those that match each covariate. Any subclass that does not contain both treated and comparison units is discarded, leaving

only subclasses containing treatment and comparison units that are exactly equal on all covariates. The benefits of exact matching are that confounding is eliminated for the included covariates and groups are perfectly balanced. However, many cases are typically discarded, which can reduce precision and limit one's ability to perform subgroup analyses. The quantity of the data across time and each racial group enabled us to use exact matching to achieve perfectly balanced comparison groups and allowed for precise measurements of the treatment effect (i.e., race) on the three federal sentencing stages (Burden et al. 2017; Stuart 2010). The covariates used for matching included age group, sex, primary sentencing guideline, district of sentencing, fiscal year, education status, whether the person is of Hispanic origin, US citizenship status, and whether the case was resolved through a trial or guilty plea. For analyses of the total sentence length and number of months below the guideline-range maximum, we included final offense level and criminal history category as matching covariates. In total, we used eight datasets for the analyses: one for each racial group for the charge and computation stages, and one for each racial group for the sentencing stages (table 2).

TABLE 2

Matched Datasets by Racial Group, Compared with the Final Sample Dataset

	Original	Matched white	Matched Black	Matched AIAN	Matched API	Matched white (sentence stage)	Matched Black (sentence stage)	Matched AIAN (sentence stage)	Matched API (sentence stage)
<b>Racial group</b>									
White	72%	60%	57%	58%	45%	61%	59%	50%	37%
Black	24%	37%	40%	20%	34%	34%	39%	11%	17%
AIAN	2%	1%	1%	20%	1%	2%	1%	38%	1%
API	2%	2%	1%	2%	19%	2%	1%	1%	44%
other	0%	0%	0%	0%	1%	1%	0%	0%	1%

Source: Authors' analysis of US Sentencing Commission individual sentencing data, FY1999 through FY2021.

Notes: AIAN = American Indian and Alaska Native, API = Asian/Pacific Islander. The total number of cases in each dataset were as follows: the original dataset = 1,281,732; the matched white dataset = 508,647; the matched Black dataset = 476,051; the matched AIAN dataset = 42,048; the matched API dataset = 50,589; the matched white dataset for sentence stage = 60,634; the matched Black dataset for sentence stage = 55,442; the matched AIAN dataset for sentence stage = 3,129; and the matched API dataset for sentence stage = 3,708.

## Dependent Measures

Racial disparities in sentencing extend far beyond just the sentences (Feldmeyer and Ulmer 2011; Fischman and Schanzenbach 2012; Tuttle 2018; Yang 2015). We therefore measured racial disparities for six elements across three stages of federal sentencing: (1) the charge stage, represented by mandatory-minimum eligibility and statutory maximum eligible sentence; (2) the computation stage, represented by the final offense level and total criminal history points; and (3) the sentence stage, represented by total sentence length and the number of months sentenced below the maximum of the guideline range (table 3).

TABLE 3

### Federal Sentencing Stages and Elements

*Six elements used for racial disparities and policy analyses*

Stage and elements	Element definitions
<b>Charge stage</b>	
Mandatory-minimum eligibility	Charged and convicted or not under a statute carrying a mandatory minimum penalty
Statutory maximum eligible sentence	Statutory maximum (in months) given all charged and convicted concurrent and consecutive maximum imprisonment months
<b>Computation stage</b>	
Final offense level	The final offense level resulting from all Chapter 2, Chapter 3, and Chapter 4 computations
Total criminal history points	The total number of criminal history points
<b>Sentence stage</b>	
Total sentence length	Total length of imprisonment including prison, alternatives of confinement, time served, and undischarged terms of imprisonment via §5G1.3 (rounded to the nearest month)
Months below guideline range maximum	The difference between the maximum possible months recommended in the guideline range and the actual total sentence

Source: Urban research team.

## Independent Measures

After creating the dataset, we created dichotomous variables for the four major racial groups the USSC collects from the presentence investigation report: (1) white, (2) Black, (3) American Indian and Alaska Native (AIAN), and (4) Asian and Pacific Islander (API) to act as the “treatment-comparison” variables. For example, belonging to the white group is the treatment group, and not belonging to the white group (i.e., nonwhite) is the white group’s comparison group. The USSC data also contain a group called “other,” which we did not examine separately because the variety of identities falling into that group

would yield unhelpful information about racial disparities across sentencing stages and over time. We did not include Hispanic origin as a separate group and instead used it as a covariate in the exact matching procedure.

## **Analytic Strategy**

We used regressions to measure the size of disparities between matched racial groups (i.e., white/nonwhite, Black/non-Black, AIAN/non-AIAN, and API/non-API) for each of the three federal sentencing stages (charge, computation, and sentence). All elements, except mandatory-minimum eligibility, were continuous variables. Thus, we employed ordinary least squares regressions where  $y$  was the stage and  $x$  was the racial group, and used the same controls used in the matching process (Nguyen et al. 2017). We took the log of each coefficient to interpret the differences as percentage differences. We did this for each fiscal year and stage and then gathered the logged coefficients for each into a time-series object. We followed the same process for mandatory-minimum eligibility; however, we employed logistic regressions and collected the exponentiated coefficients as odds ratios, for each fiscal year, into a time-series object. For consistency, we report the odds ratios as percentage differences in the odds of being convicted under a statute carrying a mandatory minimum penalty. Once we created time-series objects for each stage, we used the *strucchange* package to detect structural breaks (i.e., breakpoints) in the trend of differences for each stage across fiscal years.<sup>7</sup> Generally speaking, time-series analyses examine and predict changes over time. Measures of a topic of interest (e.g., averages) over standardized time periods (e.g., months, years, decades) are collected into a dataset called a time-series object. Regressions are used on the time-series object to measure significant ( $p < 0.05$ ) changes between time periods, which are called breakpoints. In this case, we used the measured disparities between matched groups as the measure of interest and fiscal year as the time period. We did this for each racial group and element in the sentencing process to detect significant changes in disparities over time. For the purposes of this study, we focus on the breakpoints that depict significant decreases.

## **Subgroup Analyses**

To deepen our understanding of racial disparities between stages of sentencing and over time, we conducted a series of subgroup analyses. First, we changed the reference group (i.e., white) to the other three racial groups identified in the USSC data (Black, AIAN, and API) and remeasured the disparities to examine whether the originally measured differences are consistent between demographic groups and

over time. To examine whether the measured changes differed by offense type, we compared the measured differences between white and nonwhite groups for the four most common guidelines in Chapter 2 of the federal sentencing guidelines (§2D1.1, §2L1.2, §2B1.1, and §2K2.1). §2D1.1 refers to unlawfully manufacturing, importing, exporting, or trafficking drugs; §2L1.2 refers to unlawful entering or remaining in the United States; §2B1.1 refers to larceny, embezzlement and other forms of theft, stolen property, property damage, or destruction, fraud or deceit, forgery, and counterfeiting offenses; and §2K2.1 refers to unlawful receipt, possession, or transportation of firearms or ammunition, including prohibited transactions (USSC 2018a). Consistent with the main analysis, we collected the logged coefficients and odds ratios for each subgroup into time-series objects and used the *strucchange* package to detect structured breaks across fiscal years.

## Policy Review

Using the dates when the structured breaks (breakpoints) occurred, we reviewed policies that might have contributed to significant decreases in racial disparities across the six elements of interest. Because of the stark differences between the Black/non-Black analysis compared with all the other racial groups, we included the breakpoints for the Black/non-Black analysis in the main policy review. Otherwise, we consulted the dates identified across the subgroup analyses to refine and inform our policy review. We began by searching all legislation, Supreme Court decisions, amendments to the sentencing guidelines, and US attorney general directives enacted during the year of each breakpoint and up to two years earlier. Then, two researchers well versed in federal sentencing policy and racial disparities reviewed the information and rated, on a scale from 1 (low influence) to 5 (high influence), the influence each policy likely had on the observed decrease in racial disparities for the measured breakpoint. We collected and reviewed the name of each policy, the deciding party, when it was passed and/or enacted, the goal of the policy, the population(s) it was intended to affect, and which processes of federal sentencing were changed and/or which system actors (attorneys, probation officers, or judges) were most impacted.

## Results

First, we measured the disparities between the white and nonwhite groups for the six key elements across the three stages of the federal sentencing process (see appendix A in this report's technical appendix). Then, we measured and compared disparities between matched white and nonwhite, Black and non-Black, AIAN and non-AIAN, and API and non-API groups (see appendix B in this report's



technical appendix). Then, we compared differences in the disparities for white and nonwhite groups between primary sentencing guidelines and conducted a structured breakpoint analysis to examine significant changes in racial disparities over time (appendix C). Last, we used the breakpoints during fiscal years when racial disparities significantly decreased to identify which policies likely had an impact (box 1).

---

## BOX 1

### Interpreting the Racial Disparities between Elements and Groups and Over Time

Figures 2 through 5 depict the measured racial disparities between matched cases. We report two elements for each of the three federal sentencing stages we investigated and show the size of racial disparities (in percentage difference) by fiscal year. Any disparity near or at 0 percent depicts no disparity between the racial group and matched comparison group. For all elements except mandatory-minimum eligibility, the percentage difference indicates an amount above or below their matched counterparts. A positive percentage difference indicates higher for the racial group and a negative percentage difference indicates lower. For example, in FY2010, those in the white group had an eligible statutory maximum 20 percent lower than their matched counterparts. For the mandatory minimum element, the percentage difference indicates the likelihood of being convicted of a statute carrying a mandatory minimum penalty. For example, in FY2010 the white group was 10 percent less likely than their matched counterparts to be charged and convicted with a mandatory minimum penalty.

---

### Disparities across Stages and over Time

The eligible statutory maximum sentence and a mandatory minimum penalty are dictated by the statutory ranges listed in the federal criminal penal code. The differences between white and nonwhite groups varied over time for the mandatory-minimum-eligibility element (see figure 2 on page 17). White people were less likely to be charged and convicted of a statute carrying a mandatory minimum penalty than nonwhite people, especially in the earlier fiscal years. For example, in FY2001, the white group was 34 percent less likely to be convicted of a statute carrying a mandatory minimum penalty than the nonwhite group (figure 2). But these differences decreased over time, and the white group began having an 8 to 16 percent higher likelihood of being convicted of a statute carrying a mandatory minimum penalty beginning in FY2019. The trend of differences between groups related to eligible statutory maximum sentences is similar to the mandatory minimum penalty eligibility. The disparities were largest in the earlier fiscal years (FY1999–FY2005), trending toward zero as time went on, then increasing again after FY2018. In the earlier years, the statutory maximum for the white group was

about 20 to 30 percent lower than for the nonwhite group. In the later fiscal years (FY2017–FY2018), the disparities approached zero, and in FY2019 the statutory maximum eligible sentence for the white groups increased again. By FY2021 those in the white group had statutory maximums up to 18 percent higher than those in the nonwhite group.

**TABLE 4**

**Meaning of Percentage Difference by Sentencing Element**

*Functional differences between white and nonwhite groups using FY2010 disparities*

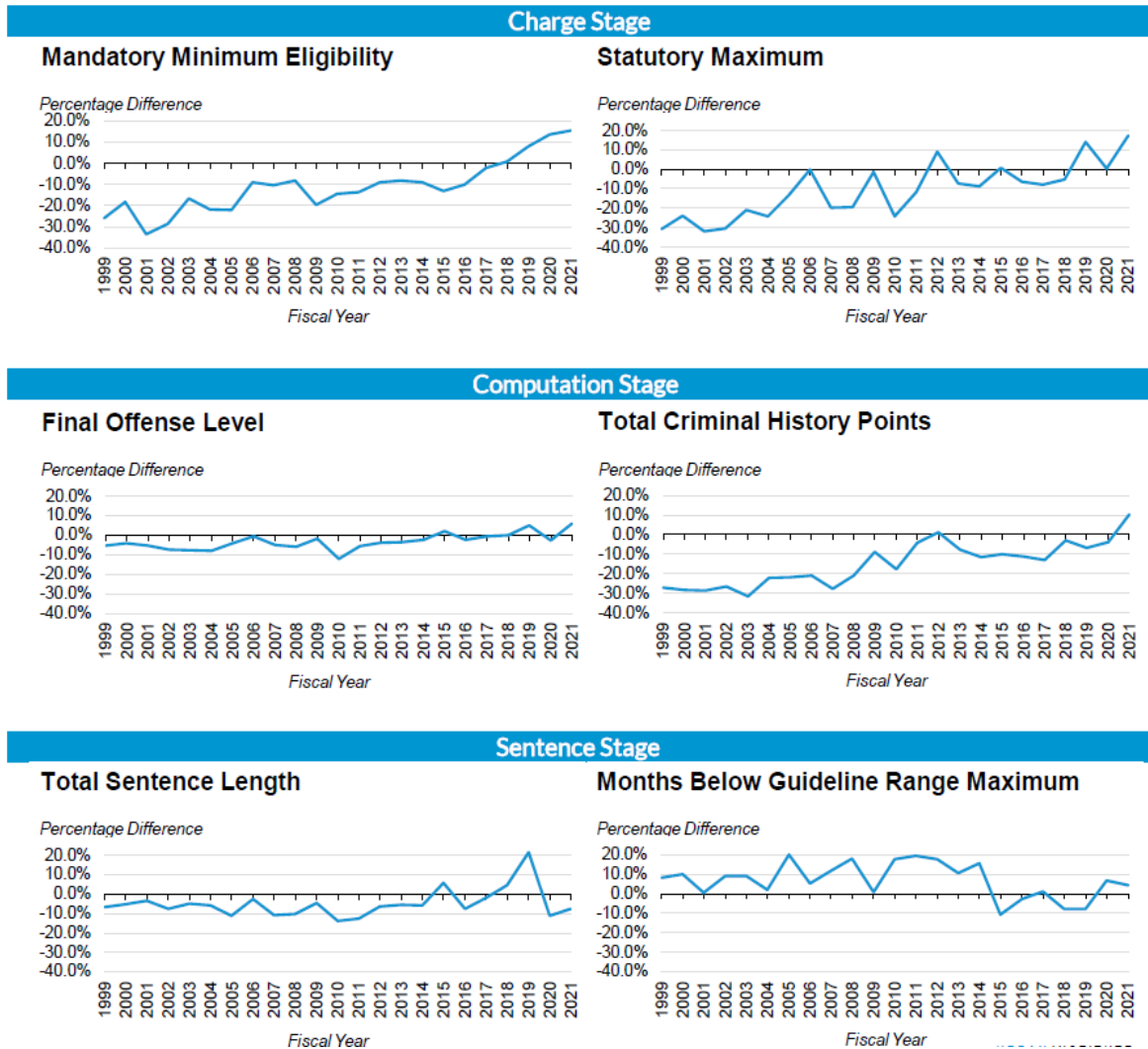
Sentencing element	Nonwhite outcome	Percentage difference	White outcome	Functional difference
Statutory maximum sentence	120 months	-24.22%	91 months	29 months or about 2.5 fewer years imprisoned
Final offense level	20	-11.97%	18	Two levels or about 10 months less in the recommended sentence range
Total criminal history points	4	-17.73%	3	One point and one criminal history level less
Total sentence length	60 months	-13.95%	52 months	8 fewer months imprisoned
Months below guideline range maximum	6 months	17.53%	7 months	1 fewer month imprisoned

**Source:** Authors’ analysis of US Sentencing Commission individual sentencing data, FY1999 through FY2021.

**Notes:** Mandatory-minimum eligibility is not included because it is a logistic regression and simply the likelihood of being convicted under a statute carrying a mandatory minimum penalty. For example, in FY2010, people who were white were 10 percent less likely to be convicted of a statute carrying a mandatory minimum penalty than their nonwhite counterparts.

We then measured the differences between the white and nonwhite groups for computation-related elements calculated by federal probation officers and used by judges during sentencing (i.e., final offense level and criminal history points). The white group consistently had final offense levels 5 to 10 percent lower than the nonwhite group (figure 2). The trend remained relatively stable until FY2014, when disparities approached zero. Conversely, the disparities between groups regarding criminal history points was much more pronounced: the white group had more than 20 percent fewer criminal history points than the nonwhite group until FY2008. Similar to the charge stage elements, the disparities between white and nonwhite groups for final offense level and criminal history points decreased over time; by FY2021, the white group had higher final offense levels and criminal history points than the nonwhite group.

**FIGURE 2**  
**Racial Disparities between the White and Nonwhite Groups**  
*Across sentencing stages for FY1999 through FY2021*



Source: Authors' analysis of US Sentencing Commission individual sentencing data, FY1999 through FY2021.

After analyzing the differences between the elements in the charge and computation stages, we used regression analyses to measure the differences in total sentence length and number of months a sentence was below the maximum end of the guideline range between the white and nonwhite groups (box 2).

---

## BOX 2

### Explaining Months below the Guideline Range Maximum

The US federal sentencing guidelines offer a sentence range based on the severity of an offense and a person's criminal history. Judges need only to consult this range and, if they do not sentence within the range, document why. In FY2011, sentences among the white group were 16 percent lower than among their matched counterparts related to the suggested maximum sentence provided by the US federal sentencing guidelines (i.e., total sentence subtracted from maximum suggested sentence). For example, a final offense level of 20 and a criminal history category of 2 results in a suggested sentence range of 37 to 46 months imprisonment. In FY2010, if the sentence of a nonwhite person was 34 months (12 months below the guideline range maximum of 46), the sentence for their white counterpart would be 20 months, which is 18 percent more months below the maximum.

---

The differences in the sentence stage elements were smaller during earlier fiscal years, increasing slightly for about five years, decreasing rapidly again, then increasing once again. For example, the white group had sentences at least 3 to 6 percent shorter than the nonwhite group from about FY1999 to about FY2005. This increased to about 7 to 14 percent shorter sentences from FY2006 through about FY2011, then rapidly decreased. Around FY2016 disparities increased again, like the other stages. Similarly, the number of months the sentences fell below the guideline maximum was 1 to 10 percent greater among the white group than the nonwhite group from FY1999 through about FY2004, then increased to about 10 to 20 percent greater from FY2005 through FY2014 before rapidly decreasing. Similar to the other stages, the disparities in total sentence length and number of months below the guideline maximum between groups eventually did not differ in the later fiscal years, with some fiscal years in which the white group had longer sentences and fewer months below the guideline maximum than the nonwhite group.

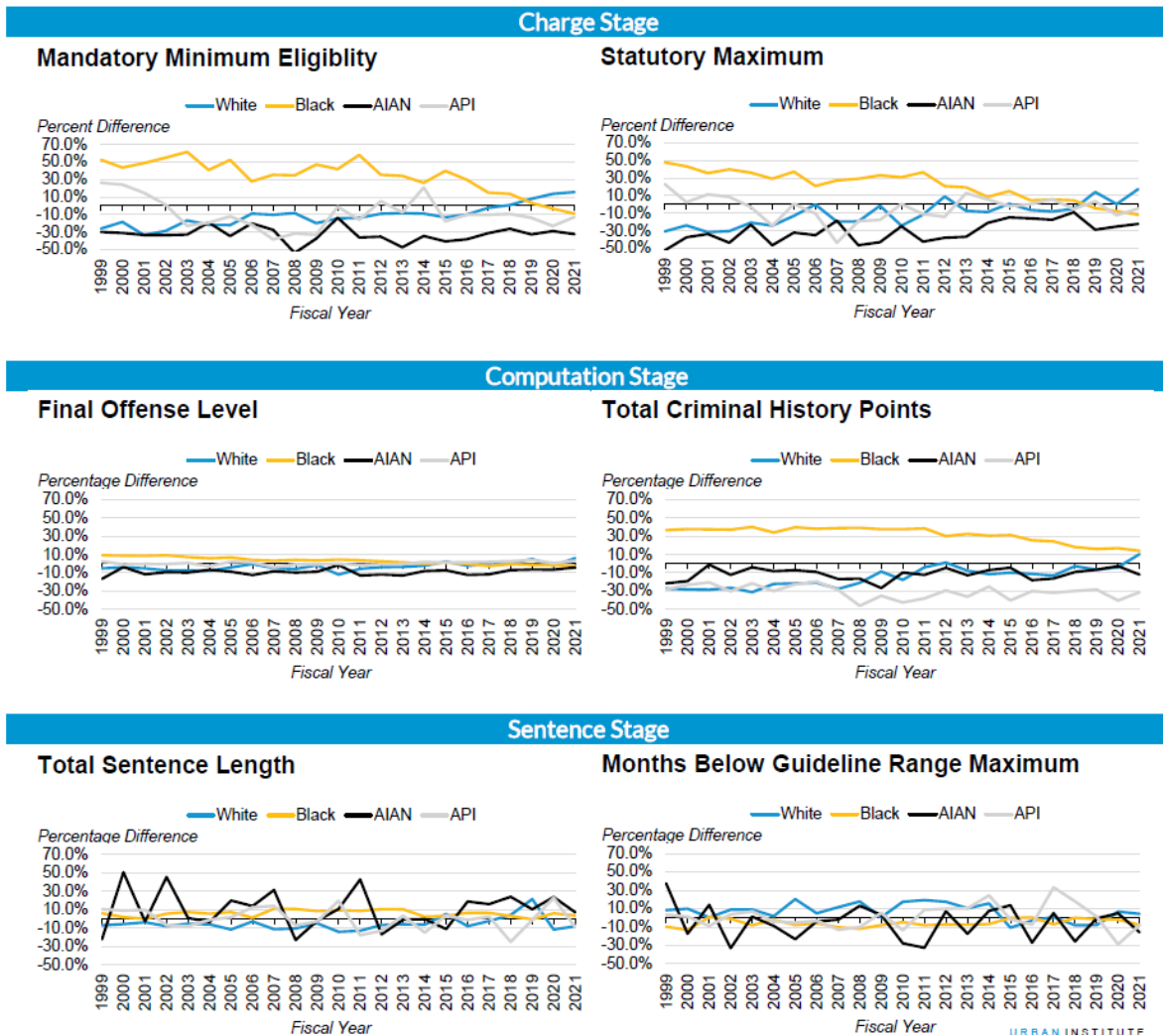
### SUBGROUP ANALYSIS: RACE

After measuring the disparities between the white and nonwhite groups, we measured disparities between the Black, AIAN, and API groups and their matched counterparts. As shown in appendix B, the size of disparities and the trends over time differed between racial groups across all three stages (figure 3).

FIGURE 3

Disparities across Racial Subgroups

Across sentencing stages for FY1999 through FY2021



Source: Authors' analysis of US Sentencing Commission individual sentencing data, FY1999 through FY2021.

Notes: AIAN = American Indian and Alaska Native, API = Asian/Pacific Islander. The race group labels indicate matched comparisons (i.e., white/nonwhite, Black/non-Black, AIAN/non-AIAN, and API/non-API).

Black people had the largest disparities and greatest disadvantages (e.g., higher eligible statutory maximums, higher criminal history scores, longer sentences) across stages and elements. Most of the measured disparities decreased over time; however, the extent of the decreases varied greatly by group, demonstrating the importance of investigating disparities among groups other than white and nonwhite, or white and Black, alone. We found the largest differences between the racial groups in the charge stage. The disparities varied over time, though the sizes and directions of the disparities differed for each group. The disparities between racial groups were most similar for the final offense level

element, in that the disparities for all racial groups were the smallest of all elements and changed similarly to each other. Most of the disparities across elements and racial groups decreased over time, except the disparities between AIAN and non-AIAN people regarding likelihood of being convicted of a statute carrying a mandatory minimum penalty, and the disparities between API and non-API people regarding total criminal history points. Those disparities changed very little over time. Also, after disparities decreased to zero, many increased again, though often with the group that previously received the advantage suddenly receiving the disadvantage.

Black people were the most disadvantaged racial group across all elements. For instance, they had higher statutory maximums, higher criminal history scores, and fewer months below guideline range maximums. Conversely, AIAN people were the most advantaged for most elements. Although those in the AIAN group typically had lower statutory maximums, were less likely to be convicted under a statute carrying a mandatory minimum penalty, had lower final offense levels, and had fewer criminal history points, they had higher sentences than their matched non-AIAN counterparts. People in the API group consistently had the lowest disparities, except in criminal history points, where they received over 20 percent fewer criminal history points than their non-API counterparts. Lastly, the disparities for the AIAN/non-AIAN and API/non-API comparisons were the most inconsistent across elements and the most variable over time, which could be due to smaller population sizes compared with the sizes of the other two race-group comparisons.

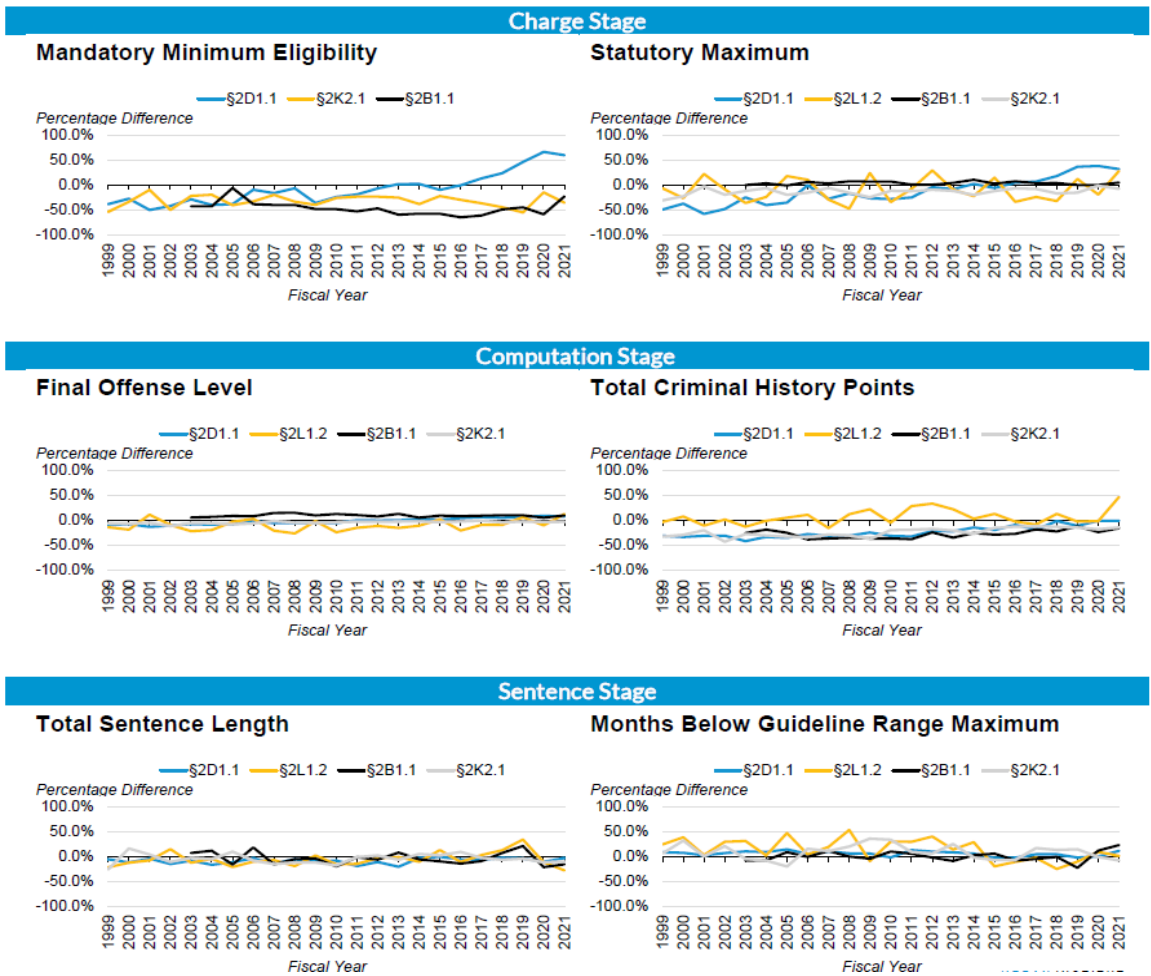
The stark differences in the disparities between racial groups, their differences between the different stages in the sentencing process, and different trends over time demonstrate the importance of measuring and accounting for disparities other than those between white and nonwhite and between white and Black people. These differences also indicate the different experiences people of different races have as they navigate the federal sentencing process, and that various policies and/or events may not impact racial disparities for every group equally.

#### **SUBGROUP ANALYSIS: PRIMARY SENTENCING GUIDELINE**

To evaluate whether racial disparities in federal sentencing vary by offense type, we conducted subgroup analyses on the four Chapter 2 guidelines most commonly used as primary sentencing guidelines (§2D1.1, §2L1.2, §2B1.1, and §2K2.1), which account for 75 percent of cases across all fiscal years. The size of the disparities between the white and nonwhite groups varied by guideline across all elements (figure 4). We found the largest differences between the US federal sentencing guidelines in the charge stage, followed by the total criminal history points element. The disparities in the charge stage elements varied most over time and had the highest disparities for all the primary sentencing

guidelines, though the patterns and directions varied by primary sentencing guideline. Similar to the main analysis and other subgroup analyses, the smallest disparities and most similar trends between primary sentencing guidelines were in the total sentence element, followed by the final offense level element.

**FIGURE 4**  
**Racial Disparities across Primary Sentencing Guidelines**  
*Across sentencing stages for FY1999 through FY2021*



**Source:** Authors' analysis of US Sentencing Commission individual sentencing data, FY1999 through FY2021.  
**Notes:** These disparities are between matched white and nonwhite cases only across primary sentencing guidelines. No statute carrying a mandatory minimum penalty exists for §2L.1.2.

For many of the elements, the disparities changed in the later years from white people having higher advantages (e.g., lower final offense levels) than their counterparts to having higher disadvantages (e.g., higher final offense levels). This change occurred with different guidelines across

different elements and at different times. Although the white group had higher final offense levels than their nonwhite counterparts in the later years, they continued to receive shorter sentences than their matched counterparts. Lastly, many of the disparities follow similar trends over time, especially between §2K2.1 and §2D1.1. The guideline that was most variable compared with the other primary sentencing guidelines was §2L1.2, which could be due to a smaller population than the other guidelines causing larger fluctuations over time, or to the nature of the offense type (i.e., unlawful entry into the United States). This is especially true for total criminal history points because federal probation officers might not be able to access international criminal history records, which could mean the criminal history calculations are incomplete.

On the one hand, differences between each primary sentencing guideline demonstrate that racial disparities vary by the offenses people are sentenced for. This makes sense given many policies target specific offenses and thus can impact racial disparities for those offenses individually. On the other hand, the data demonstrated many similarities between the primary sentencing guidelines (e.g., the size and trends of disparities for §2K2.1 and §2D1.1), indicating that policies intended for specific offenses might also impact others. Thus, it is essential to extend impact analyses beyond the apparently relevant policies, offense types, and groups to understand the breadth and depth of policy impacts fully.

## STRUCTURAL BREAK ANALYSIS

Rather than relying on our interpretation of the measured differences and our visualization of trends, we conducted a structural break analysis to detect statistically significant changes in the disparities over time. In several instances, differences between groups were not significant but changes over time were. We included these breakpoints in our policy analysis because we were focused on understanding significant changes over time, regardless of the size of disparities at those points. Although the structural break analyses detected all significant changes (i.e., increases and decreases in racial disparities), for the purposes of this study, we focused our policy review solely on the breakpoints when disparities significantly decreased (see appendix C in the technical appendix). We used these breakpoints in the policy analysis to determine which federal policies may have driven the significant decreases in measured racial disparities. Because some subgroup analyses contained small sample sizes and large standard errors, we planned to use only the dates from the white/nonwhite comparisons. But because of the extreme differences found in the Black/non-Black subgroup analyses, we also included the breakpoints for those in the policy analysis. We consulted the other subgroup analyses (race and guideline) to contextualize our policy review. The fiscal years during which breakpoints were detected are shown in table 5. The breakpoints differed between the white/nonwhite and Black/non-Black



disparities, confirming that policies likely affect multiple areas of federal sentencing at once but likely do not affect groups (e.g., people of different races) equally.

**TABLE 5**

**Fiscal Year Breakpoints Depicting Significant Decreases in Racial Disparities**

*White/nonwhite and Black/non-Black disparities, by sentencing element used for policy review*

Sentencing element	White/nonwhite disparities	Black/non-Black disparities
Mandatory-minimum eligibility	FY2005	FY2005, FY2017
Statutory maximum sentence	FY2004	FY2003, FY2011, FY2015
Final offense level	FY2014	FY2002, FY2005, FY2012
Total criminal history points	FY2008, FY2017	FY2011, FY2017
Total sentence length		FY2013
Months below guideline range maximum	FY2014	FY2014

**Source:** Authors’ analysis of US Sentencing Commission individual sentencing data, FY1999 through FY2021.

**Policy Review**

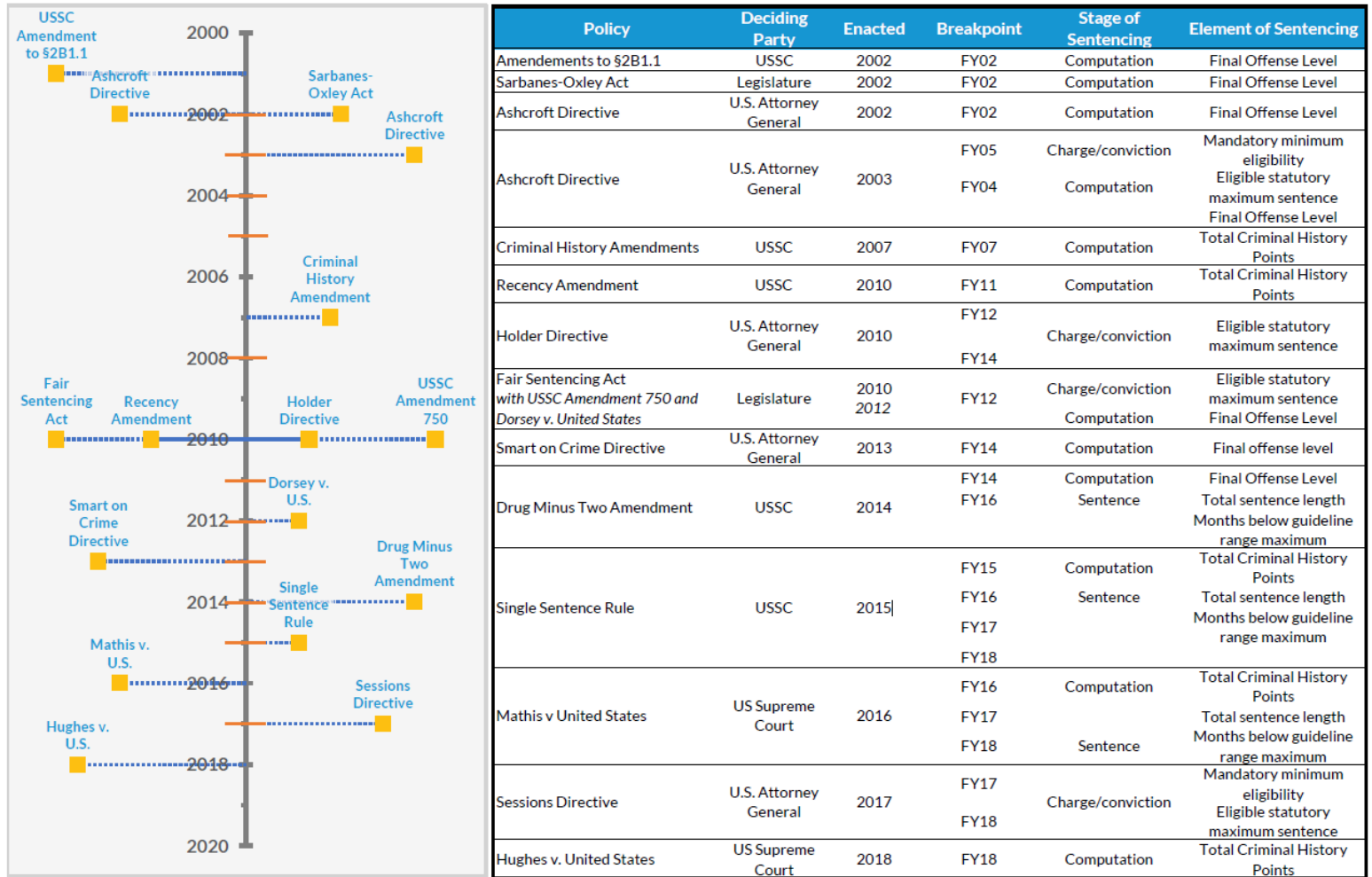
Using the breakpoints for the white and nonwhite and Black and non-Black disparities, we examined policies by each element to understand the individual influence each federal policy might have had on the elements and stages of sentencing. Afterward, we examined the identified policies across elements and stages to understand the impact of federal policies across stages and decisionmakers.

**MANDATORY-MINIMUM ELIGIBILITY**

The breakpoint in FY2005 may be related to **US Attorney General Ashcroft’s directive “Regarding Policy on Charging of Criminal Defendants”** (figure 5). That directive provided guidance to “properly enforce the Sentencing Guidelines” and instructed prosecutors to pursue the most serious, readily provable offenses. Ashcroft defined such offenses as “those that generate the most substantial sentence under the Sentencing Guidelines unless a mandatory minimum sentence or count requiring a consecutive sentence would generate a longer sentence” (Ashcroft 2003). Therefore, the more serious charges the prosecutors pursue, the more likely a person will be convicted under a statute carrying a mandatory minimum penalty. Also, this directive stated that plea bargains resulting in sentences lower than the mandatory minimums must be approved by the Department of Justice. Furthermore, charges would only be dismissed if the sentence would not be affected. This strict guidance may have led to uniformity in charging and may have therefore reduced differences in charges between similarly situated people.

FIGURE 5

Policies Likely Influencing Significant Decreases in Racial Disparities in Federal Sentencing Over Time



Source: Authors' analysis of US Sentencing Commission individual data, FY1999 through FY2021, and review of federal policies between 1998 and 2022.

Notes: USSC = US Sentencing Commission. Fiscal years with an orange line indicate dates when a significant decrease in racial disparities was measured for one or more elements.

In 2017, US Attorney General Sessions issued a directive for all federal prosecutors on department charging and sentencing policy that resembled Ashcroft's 2003 directive and that may have caused the FY2017 breakpoint. The intention of this directive was to decrease disparities in charging offenses by instructing prosecutors to again focus only on the most serious and provable offenses that would carry maximum sentences (Sessions 2017). This decision created more standardized requirements for prosecutors that may have reduced racial disparities not accounted for by prosecutors when they pursued the most serious charges and considered the whole person and case in charging decisions.

#### ELIGIBLE STATUTORY MAXIMUM SENTENCE

Because statutory minimums and maximums are set together through policies, they shared similar breakpoints and might have been impacted by similar policies during those periods. As mentioned above, **Attorney General Ashcroft's 2003 directive** limited prosecutorial discretion and instructed prosecutors to pursue the most serious charges, which are linked to higher statutory ranges. The strict guidance could have led to uniformity in charging and therefore reduced differences in charges between similarly situated people, influencing the breakpoints measured in FY2003 and FY2004.

The breakpoint in FY2011 could be explained by two major policies implemented earlier. First, **Attorney General Holder issued a directive in 2010** that encouraged prosecutors to use their discretion in making charging and plea-bargaining decisions and allowed them to individually assess cases and defendants' characteristics. The goal of this directive was to reaffirm that prosecutors should pursue the most serious charge (similar to Ashcroft's directive) but also consider the characteristics of each case and what might happen later in the criminal legal system based on the charges prosecuted (Holder 2010). Although the intention of Ashcroft's 2003 directive was to introduce uniformity of charges and reduce disparities in the application of prosecution, both of which it may have done, it did not allow for the consideration of other factors that could contribute to disparities in charges. The Holder directive allowed prosecutors to account for and try to counteract all factors contributing to racial disparities in their charging decisions. Also passed in 2010, **the Fair Sentencing Act**<sup>8</sup> reduced the sentencing disparities between crack cocaine and powder cocaine offenses from 100:1 to 18:1, which mainly impacted Black people. Although it also eliminated mandatory minimums for simple possession of crack, which lowered overall statutory ranges, there was no significant breakpoint for mandatory-minimum eligibility. This indicates the changes to possible maximum sentences were greater than the changes impacting the possible minimum sentences in this policy. And although this policy was intended for drug trafficking crimes only, the share of drug trafficking cases out of all federally sentenced cases (34 percent) could have contributed to the measured breakpoints related to the total population. The

Fair Sentencing Act might have taken several years to register a significant impact and this impact was likely spread out over time because several later policies were needed to interpret and standardize its implementation (e.g., USSC amendment 750, *Dorsey v. United States*). This may have also impacted the ability of federal decisionmakers to implement and register a change from Holder's directive.

After the FY2011 breakpoint, another occurred in FY2015, during the steady decrease in disparities between Black and non-Black people which followed that FY2011 breakpoint. The FY2015 breakpoint might have occurred because of the combined effects of the aforementioned 2010 policies and **the Smart on Crime directive that Holder issued in 2013**. That directive stated that if the quantity of the drug automatically triggers a mandatory minimum, prosecutors should not charge the quantity in cases where the defendant does not have ties to a drug organization or does not have three or more criminal history points (Holder 2013). Explicitly instructing prosecutors on how to account for criminal history or involvement in drug enterprises could have decreased the racial disparities between convicted statutes.

#### FINAL OFFENSE LEVEL

Contrary to the other elements, the breakpoints for final offense level, especially those detected from the Black/non-Black comparisons, occurred during periods of steady declines. Although disparities were already decreasing steadily, the breakpoint measured in FY2002 might have owed to the **2001 USSC amendments** to the sentencing guidelines, and subsequently to the enactment of the **Sarbanes-Oxley Act** and the accompanying **US attorney general directive** (Ashcroft 2002; USSC 2001).<sup>9</sup> The 2001 amendments to the US federal sentencing guidelines included the merging of §2F1.1 and §2B1.1 to create one main economic-crimes sentencing guideline. This merging created one of the most commonly used primary sentencing guidelines and created consistency between similar offense types in the computation of a final offense level. Soon after the 2001 amendments, Congress passed the Sarbanes-Oxley Act, which expanded and clarified several "white-collar crimes," especially those related to fraud. With the enactment of that act, Attorney General Ashcroft issued a directive clarifying how US attorneys should incorporate it into charging practices. As one of the most commonly used guidelines, these major changes to §2B1.1 and offenses it comprised may have caused the FY2002 breakpoint.

The FY2005 breakpoint related to final offense levels might be affected by similar policies that impacted mandatory-minimum eligibility and the statutory maximum because final offense levels are calculated in specific chapters of the US federal sentencing guidelines using the most serious statute of conviction and details of the convicted offenses. **Ashcroft's 2003 directive could have influenced final**

**offense levels** because it specifically emphasized the importance of uniformity in charging and would have contributed to uniformity in final offense levels for individuals with similar cases. Furthermore, prosecutors' ability to dismiss charges was more restricted because convictions needed to reflect total conduct, which federal probation officers use to calculate final offense levels. Conduct and convictions determined in the charge stage directly impact the final offense level. The requirement that prosecutors charge total conduct for everyone, rather than by their discretion, could have balanced final offense levels between groups, causing the breakpoint.

Similar to the breakpoints measured for the charge stage elements, **the breakpoint measured in FY2012 for final offense level may have been influenced by the Fair Sentencing Act of 2010**, which, as mentioned above, reduced the penalties for crack cocaine offenses. When prosecutors are charging people with drug offenses, they consider not only the type of drug but also the quantity. The larger the quantity, the more serious the charge and the higher the final offense level. Reducing the penalties for crack cocaine offenses specifically will lead to final offense levels that more closely match final offense levels of people convicted of powder cocaine offenses. It is important to acknowledge the disparities in arrests of Black people for crack cocaine versus powder cocaine; Black people are more likely to be arrested for crack cocaine whereas white people are more likely to be arrested for powder cocaine (Vagins and McCurdy 2006). This disparity leads to Black people receiving higher final offense levels because of the aforementioned extreme drug-ratio disparities. The Fair Sentencing Act adjusted this disparity for these drugs, which might be why the breakpoint occurred in the Black/non-Black comparisons only for mandatory-minimum eligibility, eligible statutory maximum, and final offense level.

Similarly, the FY2014 breakpoint might have been influenced by changes to drug policies, specifically **Attorney General Holder's Smart on Crime directive**. Like the impact on statutory maximums, by explicitly instructing prosecutors on how to account for criminal history or involvement in drug enterprises, the disparities between convicted statutes (and therefore final offense levels) were lessened. This is especially true because, in addition to statutes of conviction and drug quantity, role and previous conduct are directly tied to several base offense levels and special offense characteristics in Chapter 2 guidelines. Relatedly, in 2014 the **USSC passed the Drugs Minus Two Amendment**, which reduced the base offense levels assigned because of drug type and quantity by two levels across all drug trafficking cases (USSC 2014b).

## TOTAL CRIMINAL HISTORY POINTS

In 2007, the USSC amended the computation of criminal history directly, which may have significantly decreased racial disparities beginning in FY2008. These amendments provided clarity and specificity to federal probation officers on previously vague instructions that may have driven racial disparities (USSC 2007). For example, the USSC clarified how to account for multiple convictions named in the same charging document or that were imposed on the same day, and for convictions with concurrent and consecutive sentences. By providing explicit language on how to apply points to prior convictions, the calculation of criminal history points likely became more uniform across groups.

Similarly, in 2010, the USSC passed another amendment that affected criminal history points and could account for the breakpoint in FY2011: **the 2010 Recency Amendment**. Before that amendment, the more recent convictions were, the more criminal history points were awarded. The amendment eliminated the effect of recency on criminal history points if the offense happened within two years after release (USSC 2010). This is important to consider because when people are released from prison they are typically met with barriers when trying to reintegrate into society, such as barriers to jobs or housing, and these barriers are compounded for people from marginalized communities. As a result, people from marginalized communities return to prison sooner than their counterparts (Petersilia 2000). Black people face these barriers more so than others, resulting in longer criminal histories and more criminal history points (Finzen 2005; Wheelock 2005). The Recency Amendment could have reduced these disparities.

Several policies passed in succession from 2015 through 2018 can account for the breakpoint detected in FY2017. In 2015, the USSC passed the **Single Sentence Rule Amendment**, which, like previous amendments, clarified several aspects of previously vague language related to criminal-history-point calculations. It targeted interpretations of qualifying conduct, prior offenses, and time since prior convictions. For example, it prevented convictions that would not be counted in points from being counted as predicate offenses for Chapter 4 enhancements, including those part of a single sentence of multiple convictions (USSC 2015b). As with the other amendments, this clarification and specificity likely created uniformity between federal probation officers and cases.

## TOTAL SENTENCE

Total sentence length is directly tied to statutes of conviction, final offense level, and criminal history level. Since Supreme Court rulings such as *United States v. Booker* and *Gall v. United States*, the judge need only adhere to statutory ranges and take the final offense level and criminal history level under advisement when deciding on a sentence. But changes in racial disparities for any element preceding

the sentence stage likely will impact disparities in the sentence elements as well. The only breakpoint detected for total sentence specifically occurred in FY2013, only for Black/non-Black disparities. **Attorney General Holder’s Smart on Crime directive** likely had the largest impact, as it significantly impacted how prosecutors charged drug trafficking offenses. With that, the FY2013 breakpoint could be related to the **Fair Sentencing Act**. It likely took several years to register a significant impact, and that impact was likely spread out over time because several subsequent policies were needed to interpret and standardize its implementation (e.g., USSC amendment 750, *Dorsey v. United States*). This may have also impacted the ability to implement and register a change from Holder’s 2010 directive.

#### MONTHS BELOW THE GUIDELINE RANGE MAXIMUM

We measured one breakpoint for months below the guideline range maximum, which occurred in FY2014 for both the white/nonwhite and Black/non-Black disparities. Just like total sentence length, this element is the result of all other decisions made in the federal sentencing process. Thus, any policy passed around FY2014 likely impacted this element as well. Notably, the **Smart on Crime directive (2013) and the Drugs Minus Two Amendment (2014)** may have impacted this breakpoint by creating more standardized practices across the federal sentencing process related to drug trafficking crimes. As the most common offense sentenced federally throughout the period we studied, these changes likely drove the breakpoint despite being measured across all guidelines.

#### RETROACTIVE POLICIES

It is important to note that policies were enacted during the period we studied that applied many policies retroactively, which likely contributed, at least partially, to the breakpoints. But it was not possible for us to assess their impacts without data from before and after they were enacted. For example, the **Retroactivity of Amendment 750 (Parts A and C)** was implemented by the USSC in 2011. It applied the new drug ratio from the Fair Sentencing Act retroactively and people sentenced on crack cocaine charges received reduced sentences using the new drug ratio. Applying that act retroactively likely impacted most other elements because drug quantities dictate many charges, including those that carry mandatory minimum penalties, and final offense levels—which impact sentences and the number of months below the guideline range maximum. Also, in 2018, the Supreme Court ruled on *Hughes v. United States*,<sup>10</sup> allowing people sentenced under previous guidelines to request sentence reductions to conform to the new guidelines. Any changes this had on cases likely impacted the months below the guideline range maximum element, the total criminal history points element, the total sentence length element, and any breakpoints associated with them.

## Key Findings and Recommendations

Racial disparities differ at each key stage and for each element of federal sentencing. We consistently found the largest racial disparities for matched comparisons and the greatest differences between subgroups in the charge stage. Total criminal history points also had quite a bit of variability, high racial disparities, and large differences between groups. Conversely, total sentence length consistently had the smallest measured disparities over time and the smallest differences between groups. This is especially important because this is the element used most often in racial disparity studies. Although critical to understand, conclusions drawn about the size and nature of racial disparities in federal sentencing have likely been understated because sentence length and type have been the sole focus of such studies. With that, we only measured one breakpoint (the same breakpoint across matched racial groups) for the total sentence length element, which may mislead interpretations of policy impacts if it continues to be the only area of focus in data and policy studies. Given the differences between elements and groups, **we recommend that policymakers and researchers consider the entire federal sentencing process together when investigating disparities.**

Each breakpoint we examined indicated a fiscal year when a significant decrease in disparities between matched groups occurred. The periods with the greatest decreases in racial disparities were similar across elements and generally occurred in FY2004/FY2005, FY2012 to FY2014, and FY2016 to FY2018. Although there are similarities between the breakpoints across elements, there are differences in the breakpoints between the white/nonwhite and Black/non-Black disparities. The sentence length element is the exception, which supports the need to explore disparities in additional outcomes and groups.

To expand our knowledge beyond the traditional white/nonwhite comparisons, we measured disparities across elements and over time between matched white and nonwhite, Black and non-Black, AIAN and non-AIAN, and API and non-API people. We found that the size, direction, and trends of disparities differed greatly between groups. For example, Black people are most disadvantaged across all elements and periods (e.g., greater number of criminal history points, longer sentence lengths), whereas AIAN people had the highest advantages (e.g., lower eligible statutory maximums, lower final offense levels). Disparities were largest between Black and non-Black people, whereas disparities were consistently smallest between API and non-API people and white and nonwhite people. The experiences of people of different races throughout the federal sentencing process differ, and the breakpoints indicate that policies impact them differently. Given the stark differences in the disparities between racial groups, the differences between stages in the sentencing process, and different trends over time, **we recommend measuring and accounting for disparities among all demographic groups, in**



**addition to those between white and nonwhite or white and Black people, to understand the broad range of experiences of all people in the sentencing process.**

The differences between the disparities for each primary sentencing guideline demonstrate that racial disparities differ by the offense under which people are sentenced. Many policies target specific crimes and thus can impact racial disparities for those crimes individually. But the data demonstrated many similarities between the primary sentencing guidelines, indicating that although certain policies may not be intended for specific crimes, they may still be impacted. For example, drug trafficking offenses constitute the largest share of federal sentences and likely influenced the changes in racial disparities overall. In our policy analysis, we found that within two years of the enactment of each policy targeting drug offenses (e.g., the Fair Sentencing Act, the Crack Minus Two Amendment, and the Drugs Minus Two Amendment), overall disparities decreased significantly. But the breakpoints specific to drug trafficking offenses differed from the general breakpoints. The data show that sentencing policies related to drug trafficking clearly impacted cases other than those they were intended to impact, and their effects overlap with those of other policies. Therefore, **we recommend that policymakers and researchers extend impact analyses beyond the apparently relevant policies, offense types, and groups to fully understand the breadth and depth of policy impacts.**

Policies impact groups differently, reach across stages in the sentencing process, and impact multiple system actors. We saw this often in the charge stage, which consistently had the largest disparities over time and between groups. After each US attorney general issued a directive, a breakpoint occurred. Although each directive was intended to standardize prosecution and plea deal practices, the directives differed in approach. Each directive lacked instructions for how it was to be carried out, which likely explains the large variability in the disparities over time. Because of the apparent impact the US attorney general has on disparities, **we recommend further research into the effects the US attorney general and US attorneys have on racial disparities throughout the sentencing process.**

The importance of clear and operational policies also became apparent in our analysis. It became evident during our examination that many breakpoints were likely the combined impact of policies. Many policies identified as likely contributors to the breakpoints in disparities were enacted years before the effect was registered. During those periods, multiple policies were enacted to clarify and operationalize the original policies and likely combined to produce the registered effect. For example, the Fair Sentencing Act had a clear impact on racial disparities, particularly for Black people, but it took several years to register a significant impact. Its impact was likely spread out over time because several subsequent policies had to be passed to interpret and standardize its implementation (e.g., USSC

amendment 750, *Dorsey v. United States*). This may have also impacted the implementation of the directive issued by US Attorney General Holder in 2010. Inconsistent and vague language was the most common reason for the amendments to the US federal sentencing guidelines relating to breakpoints for total criminal history points and is likely why disparities vary so greatly over time compared with the other computation element, the final offense level. Conversely, policies such as the 2001 amendments to the US federal sentencing guidelines, followed by the enactment of the Sarbanes-Oxley Act<sup>11</sup> and the accompanying US attorney general directive, all provided detailed information on offenses and how to charge and sentence, and likely had an immediate impact beginning in FY2002. Future studies may find something similar when evaluating the impact of the First Step Act. With clear, specific, and operational language, the First Step Act was likely enacted by all system actors at all stages uniformly. In the nearly four years since its passing, no additional breakpoints have registered, and no clarifying policies have been enacted. Thus, **we recommend that policymakers include clear statements in their laws, directives, rulings, and amendments that are operational from the outset and require minimal clarification or subjective interpretation.**

Retroactive policies and social factors likely impacted the racial disparities measured in this study as well. For instance, in the later years (FY2016 and later), the disparities flipped, as those who typically experienced advantages suddenly began experiencing disadvantages, although no breakpoint registered. From contentious presidential elections, to increased conversations and awareness about racial injustices, to the later outbreak of COVID-19, any significant events in the past five years could explain the flip seen in this study. Although descriptive studies have been done on retroactive policies (e.g., Hunt, Rutter, and Kostyshak 2020; USSC 2014a, 2021), **we recommend further inquiry into the impact of retroactive policies and social factors on disparities, and additional examination of the aforementioned flip in a few years to understand the long-term change and its contributing factors.**

## Limitations

This study is new and takes a unique approach to measuring racial disparities and evaluating the effects of federal policies on them. As such, we identified several areas in which this study can be improved and expanded on. First, it is imperative to note that this study used exact matching on several qualities. Although this creates the opportunity to examine differences based on race alone, it is arguably unrealistic, because people are treated differently for many reasons, not race alone. People's experiences are affected by all qualities and environments. With that, treating race as a treatment is arguably unrealistic because race affects the development of most of the other qualities we used for

matching. Studies that explore the compounded effects of demographic characteristics across a person's life and across decisions made at each federal sentencing stage are needed. The exact matching technique also significantly reduces sample sizes, especially given the number and size of covariates, which can explain the large shifts in disparities over time for smaller groups. Although our methods are useful and informative, **research should be done with and without matching techniques that account for qualities and experiences realistically and maintain usable group sizes (e.g., coarsened exact matching, multiple regression)**. This is especially true for disparities related to citizenship and sex, which we used to match but did not explore in this study. **When possible, research should also group people according to Hispanic origin in racial disparity studies (e.g., white Hispanic, white non-Hispanic, etc.).**

We used the final offense level to investigate racial disparities. Although this provided a critical window into racial disparities at the computation stage of sentencing, it is oversimplified. The final offense level results from multiple steps in a complex process. For example, drug type and amount dictate base offense levels. The constant changes in popular drugs (e.g., the emergence of the opioid crises and influx of fentanyl) may not impact groups equally. As such, all parts of the US federal sentencing guidelines might not impact groups equally. Thus, **research should be regularly conducted to dissect the disparities for each Chapter 2 guideline, its special offense characteristics, and Chapter 3 reductions/enhancements to understand precisely which areas of the US federal sentencing guidelines drive the measured racial disparities in this study.**

The policy analysis was also simplified because of data limitations. As mentioned, we measured breakpoints for which we could not identify prominent policies that may have contributed to them. Any retroactive policy likely affected the racial disparities and breakpoints we measured. Without access to data from before and after resentencing, we cannot know the impact of such a policy. Thus, research should be done on racial disparities related to retroactive policies. We also standardized the range of dates we explored around the identified breakpoints (i.e., up to two years earlier). But confidence intervals exist in the output of structural breakpoint analyses. Thus, **research should explore the utility of using the confidence intervals of breakpoints rather than standardized ranges to identify factors that have influenced changes.**

Lastly, our analysis is limited by our having to rely on quantitative data in evaluating policy impacts on racial disparities. It is imperative to account for other perspectives that numbers cannot offer when investigating processes created and implemented by humans. Thus **researchers should pair detailed quantitative data with qualitative data.** We recommend getting policymakers' and system actors' perspectives of policies and how they are created and implemented. We also recommend getting the

perspectives of people reporting the information that constitutes this dataset to understand the lens they collect it through. And we suggest that researchers conduct case studies on all people involved in sampled cases (e.g., case file reviews) to further inform “the who” behind disparities measured in this study and others.

Racial disparities are a multilayered and multifaceted issue that no one investigation or policy will solve. It is critical to examine all data from multiple perspectives. It is essential to remember that people pass policies, implement them, and, most importantly, are punished based on them. Consistent examination and thoughtful approaches are needed to combat disparities for all people.

# Notes

- <sup>1</sup> We did not include Hispanic origin in the race comparisons because it is a separate variable in the USSC data. We focused on the race groups defined and reported by the USSC for this study. However, we recommend incorporating Hispanic origin in future research on racial/ethnic disparities by creating new race/ethnicity groups (e.g., non-Hispanic white, Hispanic-white, etc.).
- <sup>2</sup> Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, 116 Stat. 745 (2002).
- <sup>3</sup> First Step Act of 2018, Pub. L. No. 115-391, 132 Stat. 5194 (2018).
- <sup>4</sup> *Gall v. United States*, 552 U.S. 38 (2007). *United States v. Booker*, 543 U.S. 220 (2005).
- <sup>5</sup> “The Sessions Memo: Back to the Past?” Vera Institute of Justice, May 25, 2017, <https://www.vera.org/news/the-sessions-memo-back-to-the-past>.
- <sup>6</sup> For more information on exact matching and other matching approaches used via the *MatchIt* package, please refer to <https://cran.r-project.org/web/packages/MatchIt/vignettes/matching-methods.html#exact-matching-exact>.
- <sup>7</sup> Generally, structural break models detect sudden and ongoing changes in trend data. See <https://www.aptech.com/structural-breaks/> for more information. Complete information on *strucchange* is located <https://cran.r-project.org/web/packages/strucchange/strucchange.pdf>.
- <sup>8</sup> Fair Sentencing Act of 2010, Pub. L. No. 111-220, 124 Stat. 2372 (2010).
- <sup>9</sup> Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, 116 Stat. 745 (2002).
- <sup>10</sup> *Hughes v. United States*, 584 U.S. 17-155 (2018).
- <sup>11</sup> Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, 116 Stat. 745 (2002).

# References

- ACLU (American Civil Liberties Union). 2014. "Racial Disparities in Sentencing." Hearing on Reports of Racism in the Justice System of the United States submitted to Inter-American Commission on Human Rights 153<sup>rd</sup> session.
- Administration for Children and Families. 2000. *Characteristics and Financial Circumstances of TANF Recipients: October 1999-September 2000*. Washington, DC: US Department of Health and Human Services.
- Anderson, Gary Clayton. 2005. "The Conquest of Texas: Ethnic Cleansing in the Promised Land, 1820-1875." Norman, OK: University of Oklahoma Press.
- Arditti, Joyce A. 2012. "Child Trauma Within the Context of Parental Incarceration: A Family Process Perspective." *Journal of Family Theory & Review* 4 (3): 181-219. <https://doi.org/10.1111/j.1756-2589.2012.00128.x>
- Ashcroft, John. 2002. Attorney General John Ashcroft to the Director of the Federal Bureau of Investigation, memorandum. "Implementation of the Sarbanes-Oxley Act of 2002," Aug. 1, 2002. US Department of Justice Archives, Office of the Attorney General.
- . 2003. Attorney General John Ashcroft to all federal prosecutors, memorandum. "Department Policy Concerning Charging Criminal Offenses, Disposition of Charges, and Sentencing," Sept. 22, 2003. US Department of Justice Archives, Office of the Attorney General.
- Blackmon, Douglas A. 2008. *Slavery by Another Name: The Re-Enslavement of Black Americans from the Civil War to World War II*. New York: Doubleday.
- Burden, Anne, Roche, Nicolas, Miglio, Cristiana, Hillyer, Elizabeth. V., Postma, Dirkje S., Herings, Rond M., Overbeek, Jetty A., Khalid, Javaria M., van Eickels, Daniela, and Price, David B. 2017. "An evaluation of exact matching and propensity score methods as applied in a comparative effectiveness study of inhaled corticosteroids in asthma." *Pragmatic and observational research*, 8, 15-30. <https://doi.org/10.2147/POR.S122563>.
- Bushway, Shawn David. 1996. "The Impact of a Criminal History Record on Access to Legitimate Employment." PhD diss., Carnegie Mellon University. <https://www.proquest.com/docview/304235834>.
- Condry, Rachel, Anna Kotova, and Shona Minson. 2016. "Social Injustice and Collateral Damage: The Families and Children of Prisoners." In *Handbook on Prisons*, 2nd ed., edited by Yvonne Jewkes, Jamie Bennett, and Ben Crewe, 622-640. London: Routledge.
- Condry, Rachel and Shona Minson. 2020. "Conceptualizing the Effects of Imprisonment on Families: Collateral Consequences, Secondary Punishment, or Symbiotic Harms?" *Theoretical Criminology* 25 (4): 540-558. <https://doi.org/10.1177/1362480619897078>.
- Couloute, Lucius and Daniel Kopf. 2018. *Out of Prison & Out of Work: Unemployment Among Formerly Incarcerated People*. Northampton, MA: Prison Policy Initiative.
- Debrah, Ama. 2012. *Special Report: Bias Against Native Hawaiians in Hawaii Criminal Justice System*. ACLU Hawaii.
- Dietrich, Sharon M. 2002. "Criminal Records and Employment: Ex-Offenders' Thwarted Attempts to Earn a Living for Their Families." In *Every Door Closed: Barriers Facing Parents with Criminal Records*, edited by Amy E. Hirsch, Sharon M. Dietrich, Rue Landau, Peter D. Schneider, Irv Ackelsberg, Judith Bernstein-Baker, and Joseph Hohenstein, 13-26. Washington, DC: Community Legal Services, Inc. and Center for Law and Social Policy.
- Feldmeyer, Ben, and Jeffery T. Ulmer. 2011. "Racial/Ethnic Threat and Federal Sentencing." *Journal of Research in Crime and Delinquency* 48 (2): 238-70. <https://doi.org/10.1177/002242781039158>.
- Fellner, Jamie and Marc Mauer. 1998. *Losing the Vote: The Impact of Felony Disenfranchisement Laws in the United States*. New York: Human Rights Watch and the Sentencing Project.

- Ferguson, Jennifer. 2015. *Racial and Ethnic Disparities and the Relative Rate Index (RRI)*. Chicago: John D. and Catherine T. MacArthur Foundation, Safety and Justice Challenge.
- Finzen, Margaret E. 2005. "Systems of Oppression: The Collateral Consequences of Incarceration and Their Effects on Black Communities." *Georgetown Journal on Poverty Law & Policy* 12 (2): 299–324.
- Fischer, John Ryan. 2015. "Cattle Colonialism: An Environmental History of the Conquest of California and Hawai'i." Chapel Hill: University of North Carolina Press.
- Fischman, Joshua B. and Max M. Schanzenbach. 2012. "Racial Disparities Under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums." *Journal of Empirical Law Studies* 9 (4): 729–64. <https://doi.org/10.1111/j.1740-1461.2012.01266.x>.
- Franklin, Travis W., and Tri Keah S. Henry. 2020. "Racial Disparities in Federal Sentencing Outcomes: Clarifying the Role of Criminal History." *Crime & Delinquency* 66 (1): 3–32. <https://doi.org/10.1177/0011128719828353>.
- Frederick, Bruce, and Don Stemen. 2012. *The Anatomy of Discretion: An Analysis of Prosecutorial Decision Making – Summary Report*. New York: Vera Institute of Justice.
- Haley, Sarah. 2017. *No Mercy Here: Gender, Punishment, and the Making of Jim Crow Modernity*. Chapel Hill: University of North Carolina Press.
- Headley, Bernard D. 1983. "'Black on black' Crime: The Myth and the Reality." *Crime and Social Justice* 20: 50–62.
- Hinton, Elizabeth. 2016. *From the War on Poverty to the War on Crime: The Making of Mass Incarceration in America*. Cambridge, MA: Harvard University Press.
- Hinton, Elizabeth, LeShae Henderson, and Cindy Reed. 2018. *An Unjust Burden: The Disparate Treatment of Black Americans in the Criminal Justice System*. New York: Vera Institute of Justice.
- Holder, Eric. 2010. Attorney General Eric Holder to all federal prosecutors, memorandum. "Department Policy on Charging and Sentencing," May 19, 2010. US Department of Justice, Office of the Attorney General.
- . 2013. Attorney General Eric Holder to the United States Attorneys and Assistant Attorney General for the Criminal Division, memorandum. "Department Policy on Charging Mandatory Minimum Sentences and Recidivist Enhancements in Certain Drug Cases," Aug. 12, 2013. US Department of Justice, Office of the Attorney General.
- Holzer, Harry J., Steven Raphael, and Michael Stoll. 2001. "Will Employers Hire Ex-Offenders? Employer Checks, Background Checks, and their Determinants." Policy Working Paper W01-005. Berkeley: University of California, Berkeley Program on Housing and Urban Policy.
- Hunt, Kim. S., David Rutter, and Todd Kostyshak. 2020. *Retroactivity and Recidivism: The Drugs Minus Two Amendment*. Washington, DC: US Sentencing Commission.
- Johnson, Brian D. and Jacqueline G. Lee. 2013. "Racial Disparity Under Sentencing Guidelines: A Survey of Recent Research and Emerging Perspectives." *Sociology Compass* 7 (7): 503-514. <https://doi.org/10.1111/soc4.12046>.
- Kansal, Tushar. 2005. *Racial Disparity in Sentencing: A Review of the Literature*. Washington, DC: The Sentencing Project.
- Kim, Byungbae, Cassia Spohn, and E. C. Hedberg. 2015. "Federal Sentencing as a Complex Collaborative Process: Judges, Prosecutors, Judge-Prosecutor Dyads, and Disparity in Sentencing." *Criminology* 53 (4): 597–623. <https://doi.org/10.1111/1745-9125.12090>.
- Kohler-Hausmann, Julilly. 2010. "The Attila the Hun Law: New York's Rockefeller Drug Laws and the Making of a Punitive State." *The Journal of Social History*, 44 (1): 71–95.
- Kutateladze, Besiki, Vanessa Lynn, and Edward Liang. 2012. *Do Race and Ethnicity Matter in Prosecution? A Review of Empirical Studies*. New York: Vera Institute of Justice.

- LeBrón, Marisol. 2019. *Policing Life and Death: Race, Violence, and Resistance in Puerto Rico*. Oakland, CA: University of California Press.
- Le Cook, Benjamin, Thomas G. McGuire, Kari Lock, and Alan M. Zaslavsky. 2010. "Comparing Methods of Racial and Ethnic Disparities Measurement across Different Settings of Mental Health Care." *Health Services Research* 45 (3): 825–847. <https://doi.org/10.1111/j.1475-6773.2010.01100.x>.
- Mitchell, Ojmarrh, and Doris L. MacKenzie. 2004. *The Relationship between Race, Ethnicity, and Sentencing Outcomes: A Meta-Analysis of Sentencing Research*. Washington, DC: National Institute of Justice.
- Moore, Lisa D., and Amy Elkavich. 2008. "Who's Using and Who's Doing Time: Incarceration, the War on Drugs, and Public Health." *American Journal of Public Health* 98 (5): 782–786. <https://doi.org/10.2105/AJPH.2007.126284>.
- Muhammad, Khalil G. 2010. *The Condemnation of Blackness: Race, Crime, and the Making of Modern Urban America*. Cambridge, MA: Harvard University Press.
- Mustard, David B. 2001. "Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts." *The Journal of Law & Economics* 44 (1): 285–314. <https://doi.org/10.1086/320276>.
- NASEM (National Academies of Sciences, Engineering, and Medicine). 2018. *Proactive Policing: Effects on Crime and Communities*. Washington, DC: The National Academics Press. <https://doi.org/10.17226/24928>.
- Nguyen, Tri-Long, Gary S. Collins, Jessica Spence, Jean-Pierre Dures, P. J. Devereaux, Paul Landais, and Yannick Le Manach. 2017. "Double-Adjustment in Propensity Score Matching Analysis: Choosing a Threshold for Considering Residual Imbalance." *BMC Research Methodology* 17 (78). <https://doi.org/10.1186/s12874-017-0338-0>.
- Nowacki, Jeffrey S. 2016. "An Intersectional Approach to Race/Ethnicity, Sex, and Age Disparity in Federal Sentencing Outcomes: An Examination of Policy Across Time periods." *Criminology & Criminal Justice* 17 (1): 97–116. <https://doi.org/10.1177/1748895816642502>.
- Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108 (5): 937–75. <https://doi.org/10.1086/374403>.
- Petersilia, John. 2000. "When Prisoners Return to the Community: Political, Economic, and Social Consequences." *Sentencing & Corrections Issues for the 21st Century* 9 (1).
- Pinard, Michael. 2010. "Collateral Consequences of Criminal Convictions: Confronting Issues of Race and Dignity." *New York University Law Review* 85 (2): 457–534.
- Rachlinski, Jeffrey J., Sheri Lynn Johnson, Andrew J. Wistrich, and Chris Guthrie. 2007. "Does Unconscious Bias Affect Trial Judges?" Paper presented at the 2nd Annual Conference on Empirical Legal Studies, New York, July 9. <https://doi.org/10.2139/ssrn.999490>.
- Rehavi, M. Marit and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6): 1,320–354.
- Sentencing Project. 2018. *Report to the United Nations on Racial Disparities in the U.S. Criminal Justice System*. Washington, DC: The Sentencing Project.
- Sessions, Jefferson. 2017. Attorney General Jeff Sessions to all federal prosecutors, memorandum. "Department Charging and Sentencing Policy," May 10, 2017. Office of the Attorney General, U.S. Department of Justice.
- Schmitt, Glenn R., Louis Reedt, and Kevin Blackwell. 2017. *Demographic Differences in Sentencing: An Update to the 2012 Booker Report*. Washington, DC: United States Sentencing Commission.
- Schrantz, Dennis, and Jerry McElroy. 2008. *Reducing Racial Disparity in the Criminal Justice System: A manual for Practitioners and Policymakers*. Washington, DC: The Sentencing Project.
- Smith, James P. 2018. "The Long-Term Economic Impact of Criminalization in American Childhoods." *Crime & Delinquency* 65 (3): 422–444. <https://doi.org/10.1177/0011128718787514>.



- Spohn, Cassia. 2013. "Chapter 6: Racial Disparities in Prosecution, Sentencing, and Punishment." *The Oxford Handbook of Ethnicity, Crime, and Immigration*. Oxford University Press.
- Stuart Elizabeth. A. 2010. "Matching methods for causal inference: A review and a look forward." *Statistical science: a review journal of the Institute of Mathematical Statistics*, 25(1), 1–21. <https://doi.org/10.1214/09-STS313>.
- Thompson, Heather. 2019. "The Racial History of Criminal Justice in America." *Du Bois Review* 16 (1): 221–41. doi:10.1017/S1742058X19000183.
- Treuhaft, Sarah, Abbie Langston, Justin Scoggins, Joanna Lee, and Manuel Pastor, "The Racial Equity Index: A New Data Tool to Drive Local Efforts to Dismantle Structural Racism," *National Equity Atlas*, July 23, 2020, <https://nationalequityatlas.org/research/index-findings>.
- Tuttle, Cody. 2018. "Racial Discrimination in Federal Sentencing: Evidence from Drug Mandatory Minimums." College Park: University of Maryland.
- Uggen, Christopher and Jeff Manza. 2002. "Democratic Contraction? Political Consequences of Felon Disenfranchisement in the United States." *American Sociological Review* 67 (6): 777–803. <https://www.jstor.org/stable/3088970>.
- USSC (United States Sentencing Commission). 2001. "Reader-Friendly" Version of Amendments Submitted to Congress April 30, 2001 (Effective November 1, 2001). Washington, DC: US Sentencing Commission.
- . 2007. "Reader-Friendly" Version of Amendments Submitted to Congress April 30, 2007 (Effective November 1, 2007). Washington, DC: US Sentencing Commission.
- . 2010. "Reader-Friendly" Version of Amendments Submitted to Congress April 30, 2010 (Effective November 1, 2010). Washington, DC: U.S Sentencing Commission.
- . 2014a. *Final Crack Retroactivity Data Report: Fair Sentencing Act*. Washington, DC: US Sentencing Commission.
- . 2014b. "Reader-Friendly" Version of Amendments Submitted to Congress April 30, 2014 (Effective November 1, 2014)." Washington, DC: US Sentencing Commission.
- . 2015a. "Report to the Congress: Impact of the Fair Sentencing Act of 2010." Washington, DC: US Sentencing Commission.
- . 2015b. "Reader-Friendly" Version of Amendments Submitted to Congress April 30, 2015 (Effective November 1, 2015). Washington, DC: US Sentencing Commission.
- . 2017. "Mandatory Minimum Penalties for Drug Offenses in the Federal Criminal Justice System." Washington, DC: US Sentencing Commission.
- . 2018a. "United States Sentencing Commission Guidelines Manual, §3E1.1." Washington, DC: U.S. Sentencing Commission.
- . 2018b. "Mandatory Minimum Penalties for Firearms Offenses in the Federal Criminal Justice System." Washington, DC: US Sentencing Commission.
- . 2019a. "Selected Supreme Court Cases on Sentencing Issues." Washington, DC: US Sentencing Commission.
- . 2019b. "Intra-City Differences in Federal Sentencing Practices." Washington, DC: US Sentencing Commission.
- . 2020a. "Federal Sentencing: The Basics." Washington, DC: US Sentencing Commission.
- . 2020b. "Inter-District Differences in Federal Sentencing Practices." Washington, DC: US Sentencing Commission.
- . 2021. "U.S. Sentencing Commission 2014 Drug Guidelines Amendment Retroactivity Data Report". Washington, DC: US Sentencing Commission.

- Vagins, Deborah J., and Jesselyn McCurdy. 2006. "Cracks in the System: Twenty Years of the Unjust Federal Crack Cocaine Law." New York: The American Civil Liberties Union."
- Valadez, Mercedes, and Xia Wang. 2017. "Citizenship, Legal Status, and Federal Sentencing Outcomes: Examining the Moderating Effects of Age, Gender, and Race/Ethnicity." *The Sociological Quarterly* 58 (4): 670–700. <https://doi.org/10.1080/00380253.2017.1354736>.
- Wheelock, Darren. 2005. "Collateral Consequences and Racial Inequality: Felon Status Restrictions as a System of Disadvantage." *Journal of Contemporary Criminal Justice* 21 (1): 82–90. <https://doi.org/10.1177/1043986204271702>.
- Yang, Crystal S. 2015. "Free at Last? Judicial Discretion and Racial Disparities in Federal Sentencing." *Journal of Legal Studies* 44 (1): 75–111. <https://www.jstor.org/stable/10.1086/680989>.

# About the Authors

**Mari McGilton** is a research associate in the Justice Policy Center at the Urban Institute. She acts as the primary investigator and project director on federally and locally funded projects performing mixed-methods research that also employ community-engaged and participatory research methods. Her research spans topics that include cross-system involvement, racial disparities, alternative crisis response, mental health, youth and families, and community-centered approaches to safety. Before joining the Urban Institute, McGilton was a research associate for the United States Sentencing Commission, where she worked on interdisciplinary policy teams to amend the federal sentencing guidelines and inform policy decisions. McGilton previously worked for Multnomah County, Oregon, where she evaluated juvenile justice reform initiatives. As a clinician, McGilton performed neuropsychological assessments and facilitated both group and individual therapy with children and adults involved in the criminal and juvenile legal systems. McGilton received a bachelor's degree in psychology and criminal justice from Saint Louis University, and a master's degree in clinical psychology from Pacific University.

**Sabrina Rizk** is a former research intern in the Justice Policy Center, where she worked on research projects related to domestic violence, juvenile justice, prisons, and federal sentencing. She is currently a criminology master's student at the University of Maryland and holds a BA in criminology from the University of Florida. Her research interests are typically centered on racial disparities in sentencing and policing with a focus on drug policies.

## STATEMENT OF INDEPENDENCE

The Urban Institute strives to meet the highest standards of integrity and quality in its research and analyses and in the evidence-based policy recommendations offered by its researchers and experts. We believe that operating consistent with the values of independence, rigor, and transparency is essential to maintaining those standards. As an organization, the Urban Institute does not take positions on issues, but it does empower and support its experts in sharing their own evidence-based views and policy recommendations that have been shaped by scholarship. Funders do not determine our research findings or the insights and recommendations of our experts. Urban scholars and experts are expected to be objective and follow the evidence wherever it may lead.



500 L'Enfant Plaza SW  
Washington, DC 20024

[www.urban.org](http://www.urban.org)